# Agent-based Modelling of Cell-cell Interactions for *in vitro* Vascular Formation and Cancer Cell Growth

**Ye Li**

Submitted to the University of Abertay

in partial fulfilment of the requirements for

the degree of Doctor of Philosophy

November 2015

I certify that this thesis is the true and accurate version of the thesis approved by the examiners.

Signed…… Date……

(Director of Studies)

# Declaration

I, Ye Li, confirm that the work present in this thesis is my own. Where information has been derived from other sources, I can confirm that this has been indicated in the thesis.

Date……

# Agent-based Modelling of Cell-cell Interactions for *in vitro* Vascular Formation and Cancer Cell Growth

By

Ye Li

Submitted to the University of Abertay on November 17, 2015,

in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

## Abstract

A complex system is a collection of parts, that can be identical or different, that interact with each other and environment, and exhibit emergent behaviour. Here, I consider the formation of vascular structures in the body as a complex system consisting of an emergent pattern in interacting endothelial cells. A cancer tumour is a different but related complex system that contains various types of cells, some of which having cancer-inducing mutations. To understand the formation of a vascular structure or a cancer tumour, it is important to understand both the single cells and cell-cell interactions.

To study the physical interaction among cells in vascular formation or cancer cell growth, in this thesis an agent-based model is built based on the physical properties of cells which includes the size, shape, direction, and position of cells. In this way the mathematical equations in the model can show the physical variation among modelled cells. The 3-dimensional shape of cells is modelled, and so while I start with cell interactions in petri-dish the model can be easily extended to describe motility of cells in a 3-dimensional system in the future.

The physical model is implemented and then simulated with *in silico* experiments, and then the spatial distribution of cells in *in vitro* experiments is analysed and used to calibrate the model. *In vitro* experiments with and without a drug in normal and hypoxic conditions are carried out. Also the patterns formed by cells with different treatment are analysed to produce different parameter combinations in the model.

This physical model is shown to be able to predict vessel formation and be reused to predict the spatial distribution of cancer cells in *in vitro* growth experiments. With biological data such as cell size, cell shape, etc. this model is able to predict behaviours of various cell types, and can also be used to predict more complex phenomena, such as mixed type of cancer cells growing in 3-dimensions with vascular structures.

# Acknowledgements

I would like to thank my friends and family, especially my parents, for their kindness and encouragement. I also would like to thank the SIMBOS team for their valuable support : my ex-supervisor Dr. Jiujiang Zhu for instructing the derivations of physical and mathematical equations; Dr. Adam Sampson for giving suggestions about image process, parallel computing, and CoSMoS modelling approach; Dr. Yusuf Deeni and Hilal S. Khalil for giving out explanation about cell biology, carrying out the biologic experiments; Dr. Anne Savage for providing important recommendations about cell identification; Dr. Simona Hapca for helping with spatial statistics; and most importantly, my supervisor Prof. James Bown for providing a view of interdiscipline of all fields. Without their contribution, this thesis can never be finished.

# Contents

# Chapter 1 Introduction

## 1.1 Research context

Emergence is a property exhibited by a group of typically simpler entities that arises through the interactions among them, while such a property does not exist with the single entity. For example, the formation of snowflake pattern emerges as the snowflake moves through cloud of different temperature and humidity is an emergent behaviour. Swarming is another example of emergent behaviour, in which bees, fish or birds migrate together as a group. Emergent behaviour is also a topic in complex systems studies. A complex system is a collection of parts (can be identical or different) that interact with each other and environment, and exhibit emergent behaviour. Here, I consider the formation of vascular structures in the body as a complex system which consists of an emergent pattern in interacting endothelial cells. A cancer tumour is a different but related complex system that contains various types of cells, some of which having cancer-inducing mutations. To understand the formation of a vascular structure or a cancer tumour, it is important to understand both the single cells and cell-cell interactions.

In this thesis, I want to understand how the emergent behaviours in multi-cell structures are related to low-level cell interactions. Precisely, it is the relation of inter-cell physical interaction and emergent behaviour of a group of cells that I am

interested in. In particular, I hope to understand how the physical interactions amongst endothelial cells affects the pattern of vascular structure, as well as how the physical interaction among cancer cells affects the pattern of cancer cells *in vitro* but ultimately *in vivo*. The latter is particularly difficult because of cell heterogeneity within the tumour (see Section 2.1 for detail).

In the longer term, by studying the relation between cell-cell interaction and larger scale emergent behaviour in tissue structures, it is able to understand how tissue or organs grow, which is useful in treatment of tissues and organs. In particular, if how vascular structure is formed is understood, then it is possible to understand the factors that affect the growth of vascular structure. Similarly, the study of relation of cancer cell interaction and tissue-level behaviour enables people to understand how tumour is formed and what affect the growth of tumour. At last such knowledge suggests methods to control the growth of vascular structure or cancer tumour.

Furthermore, as cancer treatment drug acts differently on different cell types and different individual cells, it is difficult to predict the efficiency of drugs on controlling the growth of cancer tumour. However, studies on the effect of drugs on low-level interactions among cancer cells, may inform drug design. Note that in this thesis, as an early step to tissue growth I am limited to and so focus on *in vitro* growth in petri-dish, in which 2-dimension patterns are formed; likewise ‘cell interaction’ means the cell interaction in petri-dish. It is recognised that there are two important issues of

this limit. One is the cell behaviour in *in vitro* condition is not completely the same as with *in vivo*. Thus there is an underlying assumption here, that the driving mechanisms of cells interaction are same (or sufficiently similar) in both *in vitro* and *in vivo* systems. In addition, the *in vitro* system gives experimental control, which is traded off against the biological reality *in vivo*. The other issue is: cells grow to 2-dimensional structures (monolayers) in petri-dish, while *in vivo* cell growth forms 3-dimensional structures. In other studies, *in vitro* 3-dimensional experiments have been introduced, in which cells form 3-dimensional structures (spheroids). Based on the experiments in petri-dish, spheroids can help as a bridge to connect 2-dimensional in vitro experiments with 3-dimensional ones. The model developed in this thesis can be extended to represent such 3D structures.

To study the physical interaction among cells in vascular formation or cancer cell growth, I follow an agent-based modelling approach. First, I will find a method to represent the physical properties of cells by mathematical equations, which should include the size, shape, direction, and position of cells. In this way the equations can show the physical variation among modelled cells. I will start with cell interactions in petri-dish, thus the modelled cells move on substrate plane. As I will model the 3-dimension shape of cells, it will be a 3D model describing 2D cell motility. Later it can be easily extended to describe motility of cells in 3-dimensional system.

Currently, the modelled cells move on substrate plane, thus the physical interactions

happen among nearby modelled cells, and space is important. With mathematical equations of physical properties of cells, I can build equations to describe the spatial relation between a pair of arbitrary modelled cells. Here I introduce the concept of contact potential, see Section 2.4.4 for detail. After that, the physical interactions among cells are modelled with a series of equations relative with the contact potential.

Using existing results of vascular formation experiments *in vitro*, and the spatial feature of vascular network will be analysed to calibrate the physical model. It was not possible to carry out the necessary experiments to calibrate the physical model of vascular formation, and so I converted the model to a new system where I had data streams available to me.

Similar to the vascular formation model, I hope to understand the nature of the interactions among cancer cells by analysis of the patterns that they form. In *in vitro* experiments, the physical interactions among cancer cells affect the spatial feature of the pattern formed by cells. Thus I will keep track of cell position and carry out spatial analysis on positional data. I also hope to understand the effect of a growth inhibitor drug and environmental effects, specifically hypoxia (see Chapter 5), on the physical interactions among cancer cells. Further, I are interested in the different responses of different cell lines that are sensitive and that are resistant to the drug in order to explore the dynamics of heterogeneous populations. Therefore independent experiments will be carried out for the each cell line, with and without drug in normal

and hypoxic (oxygen starved) conditions.

The physical model will be simulated with *in silico* experiments, and then the spatial distribution of cells in *in vitro* experiments will be analysed and used to calibrate the model. I will compare the pattern formed by cells with and without drug in normal and hypoxic conditions. Experiment with different treatments may require different parameter combinations in the model. I consider the differences in parameter sets when modelling the effect of drug action hypoxia and their combination.

The method of comparing *in vitro* and *in silico* experimental results is based on a set of hypotheses. See Section 1.2 for detail.

## 1.2 Hypothesis

The aim of this thesis was to develop an individual-based model of populations cells and inter-cell interactions. The model focused on characterising the physical properties of cells and how these drive inter-cell interaction dynamics. The biological behaviour of cells was captured in terms of changes to the physical property of cells, such as cell size and shape, and assumes that the physical model can be used to predict biological phenomenon. This assumption led to the following hypotheses:

1 An in silico physical model can reproduce important behaviours of human cells in experimental systems;

2 The model can be parameterised with biologically interpretable values to model differences arising from experimental conditions for the same biological system;

3 The characterisation of cell distributions by spatial statistics enables prediction of biological phenomena using the physical model.

## 1.3 Thesis structure

This thesis is structured as follows:

Chapter 2 covers reviews of related studies about cancer cell and modelling approach. In Section 2.1, cancer is described as a complex system with six hallmarks. Then two main types of complex system modelling approaches are reviewed in Section 2.2. Also I review the CoSMoS modelling approach which concepts are used for model repurposing from vascular formation to cancer cell growth. After that, three modelling approaches, each showing different aspect of cancer research, are detailed reviewed in Section 2.3 as case studies; and all leads to the conclusion that physical interaction plays an important role in tumour morphology. Considering the type of modelling approach, a list of issues in modelling physical interaction of cancer cell are discussed in Section 2.4, and for each issue a solution is found and conclude in the end of Chapter 2. In addition, in Section 2.5, the possible method to validate and calibrate the spatial distribution of cancer cell is discussed.

Following Section 2.4 and to address Hypothesis 1 and 2, a detailed physical interaction model is built in Chapter 3. The shape of cells is modelled in Section 3.1. With shape and position of two modelled cells, the contact potential and adhesion potential are estimated in Section 3.2 and 3.3 separately. In Section 3.3 the adhesion force is calculated from adhesion potential. From the contact potential, the energy is calculated by Hertz model; and from energy the contact force is calculated. This approach can be found in Section 3.4. Section 3.5 shows the calculation of contact and adhesion torque, which controls the rotation of modelled cell. The velocity and angular velocity of modelled cells are calculated from force and torque in Section 3.6; and at last the velocity and angular velocity are used to calculate the position of modelled cells in Section 3.7. Note that as an agent-based model, the set of formulas with which to calculate all the physical quantities are the same for every modelled cell, while the actual value of quantities may vary. In section 3.8, the physical interaction between a modelled cell and the substrate is simplified as the physical interaction between the same modelled cell and its mirror image of substrate plane.

Using the physical model built in Chapter 3 and following the computational implementation approach discussed in Section 2.5, the physical model is first used to predict the early stage of vascular formation in Section 4.1. Although I do not have *in vitro* experimental data to calibrate the model for vascular formation, I are able to present result of *in silico* simulation with different number density of modelled cells

in Section 4.1.3, with which the Hypothesis 1 is partly confirmed. In section 4.2.2, the reusability of the physical model, including additional requirements of modelling the cell cycle and population growth, is briefly summarized (drawing on a published first author journal paper). To explore Hypothesis 1 and 3, a set of eight experiments are designed afterwards (in Section 4.2.3).

Following the reusability analysis in Section 4.2.2, the cell cycle and cell population growth are modelled in Section 5.1 and 5.2 separately; and the result of experiments designed in Section 4.2.3 is shown in Section 5.3. Note that the experiments are expected to be used to calibrate both cell cycle/population growth and spatial distribution; the experimental result about spatial distribution is shown in Section 6.1. Section 5.5 demonstrates the process to calibrate the model with experimental result, which confirms the Hypothesis 1 and 2.

With population growth curve calibrated in Chapter 5, experiments (also designed in Section 4.2.3) are carried to calibrate spatial distribution of modelled cells in Chapter 6 (this also confirms Hypothesis 3). The experimental results are images taken from experiments. The images are firstly passed through a pre-analysis process to enhance the quality of images in Section 6.1.1. Then in Section 6.1.2 and 6.1.3, the positions of cells are extracted from processed images. In Section 6.2, the position data is analysed by various spatial statistic tools to provide basic understanding of characteristic of patterns formed by cancer cells in experiments, from which I find that the data of

control group does not reflect same number density of cells with Chapter 5. The control group is then removed from Section 6.3, in which spatial data is used to calibrate the model.

Chapter 7 concludes the three hypotheses and discusses potential next steps of future work.

# Chapter 2 Literature Review

## Introduction

In Chapter 1 cancer is considered as a complex system. Thus in this chapter, i firstly detailed review cancer and vessel formation as complex systems, then go through modelling approaches of complex systems in Section 2.2. In Section 2.2.5, three models about various aspects of cancer are reviewed as case studies. As I aim to study the emergent behaviour of physical interaction among cells, equations will be built based on physical properties of cells discussed in Section 2.3.1. The software development approach is discussed in Section 2.4, with a computation implementation to speed up *in silico* experiment. Finally the method of model calibration is discussed in Section 2.5.

## 2.1 Cancer and vessel formation as complex systems

"A complex system is any system featuring a large number of interacting components (agents, processes, etc.) whose aggregate activity is nonlinear (not derivable from the summations of the activity of individual components) and typically exhibits hierarchical self-organization under selective pressures." (Indiana University Bloomington website http://www.informatics.indiana.edu/rocha/complex/csm.html) In other words, in complex system, large number of (relatively simple) objects interact and show a new behaviour as a whole. Examples of complex systems are: social

systems (formed of people), the brain (formed neuronal cells), molecules (formed of atoms), and the weather (formed of air flows) (New England Complex Systems Institute (NECSI) website http://necsi.edu/guide/). In this thesis, cancer (formed of cancer cells, base tissue and vascular) and the vessel formation process (performed by endothelia cells) are referred as complex systems, with their features that suits description of the complex system explored in the following paragraphs.

# Cancer

Cancer is a class of diseases where mutated cells can over-grow and over-reproduce (Bown J., *et al*. 2015). There are six biological hallmarks of cancer, including sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis; and these capabilities enable tumour growth and spreading, termed metastasis (Hanahan *et al*, 2000. Luo *et al*, 2009. Hanahan *et al*, 2011).

The normal cells are controlled by proliferation and growth signals that ensure the normal cells enter and progress through their natural development according to a proper cell cycle. The mutated cell, however, has the capability to sustain proliferative signalling in alternative ways (Hanahan *et al*, 2011). See Goltsov *et al* (2010) for more on mutations and aberrant

One such observed mutation in cancer cells is in the RB pathway, which decides

whether a cell should proceed through its cell cycle in the normal manner. Mutations in this pathway can lead to abnormal progression through the cell cycle, and in cancer this can mean excessive proliferation. The TP53 protein can halt cell cycle in response to improper growth conditions (hypoxia, lack of space or nutrient), or trigger programmed cell death in response of cell damage (Toettcher *et al*, 2009). In these cancer cells deregulation of this RB pathway or the TP53 protein can thus evade normal growth suppressors (Nevins 2001. Vousden *et al*, 2009. Muller *et al*, 2013).

A mutation in TP53 also grants cancer cells resistance to cell death. The pre-programmed cell death works as a barrier to cancer development (Elmore, 2007. Ouyang *et al*, 2012). Researches revealed that the apoptosis (death) mechanism is weakened in tumours high with malignancy and resistance to therapy (Hanahan *et al*, 2011). Apart from loss of TP53 function, cancer cells have a range of other strategies to evade the apoptosis mechanism. Tumours may increase expression of anti-apoptotic regulators (Bcl-2, Bcl- $x_L$ ) (Kang *et al*, 2009. Kelly *et al*, 2011. Thomas *et al*, 2013.) or of survival signals (lgf1/2) (Samani *et al*, 2007. Huang *et al*, 2010.). Tumours may also "short-circuit" the extrinsic ligand-induced death pathway (Hanahan *et al*, 2011).

Another irregular behaviour of the cancer cell is replicative immortality and this is driven by changes in the function of telomeres, structures at the end of chromosomes. The telomeres, which consist of repeated hexanucleotides (Sadava *et al*, 2011), are involved in the capability of the replicative immortality (Bodnar *et al*, 1998. Hanahan

*et al*, 2011. Yaswen *et al*, 2015). In normal cells, the telomeres shorten following cell proliferation and when the telomeres get too short the cell can no longer proliferate (Passarge 2007). In this way control the number of generations the normal cells can pass through. The cancer cells, on the other hand, maintain the length of telomeres by over-expression of telomerase, and lead to immortal cells (Aschacher *et al*, 2015).

On a larger scale, the tumour tissue is made up of various types of cells, including cancer cells of different types (Bown J., *et al*. 2015). The tumour tissue has another 2 hallmarks of cancer. Unlike normal tissue, the vascular network, delivering the blood necessary for rapid cell growth, is typically activated and pervasive in tumour tissue. The blood vessels in tumour tissue are often enlarged and mishapen (Jain *et al*. 2005. Goel *et al*. 2012). The tumour also has the capability of invasion and metastasis, which is in part the result of downregulation (decrease of quantity of a cellular component such as RNA or protein in response to external stimuli) of the cell-to-cell adhesive molecules enabling cell movement (Yamaguchi *et al*. 2005. Ramis-Conde *et al*. 2009. Friedl *et al*. 2011. Canel *et al*. 2013). In addition, the tumour tissue can produce a microenvironment to evade immune destruction (Bown J., *et al*. 2015).

In summary, cancer is a complex system comprising various types of normal and cancerous cells, which show irregular and complex individual and emergent behaviour on cellular and tissue scale. A key issue in cancer cell biology is to understand the relationship between dysregulation at the cellular level and tumour

formation and metastasis at the tissue and body level. This is necessary to understand because anti-cancer drugs operate at the cellular level and seek to control behaviour at the system scale to affect tissue structures.

## Vessel formation

The blood vessel is formed by endothelia cells. The phenomenon of vascular formation contains two main mechanisms: vasculogenesis, in which dispersive endothelial cells form vascular network; and angiogenesis, in which new vascular sprouts from an existing vascular network. In all the following parts of thesis, the phrase ‘vessel formation’ and ‘vascular formation’ refer to vasculogenesis.

As stated in (Scianna *et al*. 2013), many *in vitro* vessel formation experiments have been undertaken with various endothelia cell lines and substrata, from which the overall process of vasculogenesis is observed and reported. The process contains following phases:

(1) The dispersive endothelial cells move in the direction of nearby regions with high cell densities (Serini *et al*. 2003; Ko¨hn-Luque *et al*. 2011). It takes around 3 to 6 hours until cells contact their neighbours. Cell migration contains a small random component. Also in this phase the speed of cells is higher than other phases;

(2) After cells make contact with their nearest neighbours, they elongate their shapes,

and the sites connecting cells and the substrate increase. Eventually the endothelial cells form a network; (Scianna *et al*. 2013)

(3) The multi-cell network slowly moves as a whole and goes through a thinning process. During this phase the structure does not deform much; (Scianna *et al*. 2013)

(4) In the last phase every single endothelial cell folds up and form the lumen of the vessel. (Kubota *et al*. 1988)

The total process takes 9-15 hours (Scianna *et al*. 2013). In some modelling studies the above phase 3 and 4 are considered as one (Serini *et al*. 2003). In this thesis I focus on phase 1, where the shape of cell does not change, and this allows me to consider the motility of the cell and the parameters that affect it. In phase 1, the growth factor VEGF (vascular endothelial growth factor) is considered one of key factors of vasculogenesis and angiogenesis (Karamysheva, 2008). The concentration of VEGF controls the direction of cell motility (Serini *et al*. 2003). In (Gamba *et al*. 2003), it has been established that the typical size of the network is proportional to the diffusion coefficient and half-life of VEGF-A. Further, in the in vitro experiment in (Serini *et al*. 2003), eliminating the effect of VEGF-A (by adding antibody of VEGF or removing the VEGF-A receptor) prevents the vessel formation.

In summary, in the early stage of vessel formation, large number of endothelia cells

interact with each other and environments, and the result of the interaction is the net-shaped vessel structure, which is considered self-orgnisation. Each endothelia cell is driven by the same rule, but a group of cells show different behaviour (can form the vessel). Therefore it is considered as a complex system in this thesis.

## 2.2 Complex systems modelling process

### 2.2.1 Modelling process generally

As discussed in section 2.1, cancer is a complex system, which makes it a challenging to understand and then manage. I need methodologies to systematically understand the key factors in cancer formation, growth, invasion and metastasis, as well as how those key factors work together. Systems models, coupled to experimental systems, can describe such interactions amongst the different factors of the system. Traditional science studies a system by dividing the system to components, and studying each component individually, i.e. a reductionist approach. However, this approach presents several difficulties for developing an understanding of the cancer system. First, not all the cells in the cancer are identical. Second as most of cells are fixed in the stroma, the interaction among cells often occurs between neighbours, thus the spatial arrangement of cells must be considered. Following on from these, since cancer is so complex no single experimental approach can reflect all its characteristics; models must often base on data sets from different experiments. Finally, cancer presents the hallmarks mentioned above on both tissue scale and cell scale, the connection

between cell scale behaviour and tissue scale behaviour is hard to measure (Bown J., *et al*. 2015). Broadly, the modelling approaches for cancer can be divided to two types: data-driven and process-based (Bown J., *et al*. 2015).

## 2.2.2 Data-driven modelling

Data-driven modelling approaches normally focus on data sets arising from experiments and study the connections between them, for example correlations, regardless of the organising principles and underlying processes that drive the system. Here, for illustration two such data-driven models are outlined: one at the tissue scale and one at the cellular scale.

Savage *et al*. (2013) provides a data-driven modelling approach connecting the topologogical features of tumours with patient prognosis by regression analysis. In this study the Minkowski functionals: volume (V), surface area (SA), integral mean curvature (IMC) and integral total curvature (ITC) are chosen to form quantitive measurements of geometrical and topologogical features of 3-dimensional images of tumour. The main finding was that different tumour subtypes of the HER2 status show different responses to the anti-cancer drug tamoxifen in terms of the Minkowski functionals measurements. Additionally, the influence of four explanatory variables: tamoxifen treatment, ER status, HER2 status and node status on three morphological measurements: IMC/V, ITC/V and SA/V are analysed by linear regression. Several observations are made based on the results: the volume of both ER+ and ER- tumours

is changed in response to tamoxifen treatment; tumour with differing HER2 status produce similar morphological measurements without tamoxifen treatment; HER2- tumours show a response to tamoxifen treatment in morphological measurements while HER2+ tumours do not. Finally the tumour grade is assessed by discriminant factor analysis with Minkowski functionals as independent variables and the result shows near-perfect accuracy.

In a large, data rich study of cellular responses to sequential drug treatments; Lee *et al*. (2012) explore the relation between patterns of drug administration and measures of levels of activity of key signalling proteins. These measures are used to define cell-type differences and drug-treatment specific differences among cell lines and for different treatment regimes by principal component analysis and partial least squares regression. To assess the synergistic actions of drugs, several drug combinations were tested in different breast cancer cells with different order and timing of drug addition. Then it was found that addition of erlotinib at least 4 hours prior to doxorubicin causes significantly enhanced apoptotic response. As the effect of erlotinib is EGFR suppression, then several experiments were designed to test if the increased apoptosis is caused by EGFR suppression. Then differentially expressed genes (DEGs) are measured on BT-20 cells with 30 minutes, 6 hours and 24 hours of erlotinib treatment. It shows that the number of DEGs is dramatically increased in the groups of 6 hours and 24 hours of treatment. Comparing with BT-20 cells, number of DEGs with 24 hours treatment of HER2 cells and MDA-MB-453 cells is not significantly changed.

After that, a model is built using principal component analysis (PCA) and partial least squares (PLS). From the model 4 proteins as signals are identified: cleaved caspase-8, cleaved caspase-6, phosphor-DAPK1, and phosphor-H2AX. Then the 'variable importance in the projection' (VIP) is calculated for each signal and caspase-8 is identified as the major signal. Thus caspase-8 is hypothesized as main factor of signalling pathway that increases the sensitivity of BT-30 cells to doxorubicin, which are proofed in *in silico* test.

The two exemplar case studies show that the data-driven modelling can build models on both cell and tissue scale. As data-driven modelling makes use of existing statistical model to process data sets, it is relatively easy, compared with process-based models, to build a data-driven model and explain its output. Unlike the process-based modelling approach, data-driven modelling does not require extra assumptions about unknown parts of the system. However the relation between data set to build data-driven model and the data produced by the model needs to be carefully examined. If the two sets of data are similar in nature, the data-driven modelling can provide high accuracy. Otherwise the data-driven modelling cannot provide valuable conclusion.

### 2.2.3 Process-based modelling

Process-based modelling, on the other hand, focuses on the key elements within the system and the way the key elements and their interactions affect the behaviour of the

system. Process-based modelling includes mathematical modelling such as ordinary differential equation models and computational approaches such as individual-based models. Goltsov *et al*. (2013) modelled the sensitivity of an intra-cellular signalling network. In this study an ODE-based model is used to study the mechanisms of resistance in RAF/MEK/ERK and PI3K/PTEN/AKT signalling network in PE04 cells. Then the model is calibrated by the experimental data. From simulation of the model and experimental data, it is shown that PTEN lose or PI3CA mutation can cause cancer cells to gain resistance to drug. More tissue scaled process-based modelling approaches are reviewed in section 2.3.

There are three main types of process-based modelling: continuum, agent-based, and hybrid. The difference among these modelling types is to consider cell as continuous medium, as a particle, or as a combination of both.

Specially, individual-based modelling, or agent-based modelling, defines a set of rules that act on every agent, and the computational result of a (typically) large number of agents has the ability to show behaviours on emergent scales, for example cell-cell interactions showing emergent behaviour at the tissue scale. In current cancer biology both cell signalling and tissue morphology are seen as very important (D. Vaudry *et al*. 2002; Liming Xiong *et al*. 2002; Zhengyu Ma *et al*. 2002; Markku Miettinen and Jerzy Lasota, 2006). However, the relationship between individual cell behaviour and the tissue morphology, i.e. the tumour cell and tumour tissue, and the endothelial cell

and vessel, etc., is not only hard to observe but also hard to measure. The complexity of tissue-like tumours presents the challenge of connecting molecular control mechanism with the tissue-based behaviour (Pamela K. Kreeger and Douglas A. Lauffenburger, 2009).

Typical cell-based models in vessel formation studies (Roeland M.H. Merks *et al*. 2006) are lattice-based models that follow the energy minimization assumption. The energy minimization concept is described in section 2.2.1. The calculation of energy is a Metropolis method, as discussed in 2.3.1. Thus instead of inter-cell dynamics being controlled by concentration of chemotaxins alone, the motility of endothelia cells also tends to reduce the total energy of the whole system. The cell-based models also produce results well agreed with *in vitro* experiments. Unlike continuum models, cell-based models are able to include a mechanism of intra-cellular signalling pathways.

The study (Alvaro Ko¨hn-Luque *et al*. 2011) assumes that the VEGF is not autocrine, and add a VEGF-relevant signalling pathway to the model, enabling this model to reproduce early *in vivo* vascular assembly process. The inclusion of modelling pathways also enables cell-based models to reproduce malignant vasculogenesis by adding other irregularities in signalling pathways, i.e. mutations. Models such as (Marco Scianna *et al*. 2013) are used to study vessel formation in tumours.

The continuum model, on the other side, is widely used to study the motion of tissue as a whole. For instance, in (King J R and Franks S J, 2004), it introduced equations to describe tissue growth. In all these models cells are considered as continuous medium that evenly distributed within tissue. Then the characters of tissue and interaction between tissue and environment, which are normally described by a set of partial differential equations, are studied. The continuum model can be used to represent multi-type cells, however it is restricted that within a tissue the cells must be homogeneous, and cell colonies of different types are considered as different tissues. Also it is easy to track the physical and biological status of each modelled agent.

The continuum models assume the endothelial cells are continuously distributed on the substrate and thus use partial differential equations to represent the rules that affect motility of endothelial cells, e.g. the momentum balance of the cells, the distribution of chemotaxins, etc. Early studies assume that the pulling action of endothelial cells is the main driving mechanism and therefore three partial differential equations are derived, representing each of conservation of cell density, extracellular matrix density and pulling force balance (Namy *et al*. 2003). After that, by observing *in vitro* experimental processes, the effect of chemotaxins is added into the equations (Serini *et al*, 2003), (Gamba *et al*. 2003) and (Ambrosi and Gamba *et al*. 2004). Later model studies paid more attention to the chemotaxins - precisely chemoattractant - and its effect on cell motility. The direction of cell movement, motility persistence and friction between the moving cells and substrate are modelled. There are also

additional assumptions: the endothelial cells do not divide or die during the process (Neufeld *et al*. 1999); connected cells do not overlap; and the release of chemoattractant is faster than its degradation.

The *in silico* simulation reveals the driving force of chemoattractant, supporting the assumption of the driving role of VEGF-A in the first phase of vessel formation (Serini *et al*, 2003). Finally the most recent modelling approaches consider the effect of more than one type of chemotaxins. Both chemoattractant and chemorepellent are considered in vessel formation (Lanza *et al*. 2006), and the *in silico* result shows that the chemorepellent is more effective in driving vascular formation.

As discussed in the Introduction to this chapter, the purpose of this modelling approach is to study how the rules that controls single cell may affect the emergent behaviour of group of cells.

In general, agent-based modeling (Bown J. *et al*. 2012) may be a suitable tool to understand the cell-scale behavior between the inter-cell level and organ level dynamics; the shortage of direct experiment data, no matter in vitro or in vivo, precludes clear verification. The signaling-scale study itself is subject to lack of data for model construction and validation (Faratian, Dana *et al*. 2009). The tissue-scale data, on the other hand, provides possible indirect support. Unfortunately it has its own disadvantages. As mostly this kind of data is collected at certain time points, it

cannot display the dynamic of tissue behavior (Gabriele Bergers and Douglas Hanahan. 2008), which is essential in cell-scale interaction study. For example, in the early development of tumor in vitro experiment, there is no vessel at first, the angiogenesis happens following necrosis of the center of tumor (Serini *et al*. 2009).

As a process-based modelling approach, agent-based modelling has its limits. Firstly comparing to data-driven modelling approach, the process-based modelling often requires assumptions on unknown part of the system, which may introduce errors. At the same time in the process-based modelling it is hard to identify errors in model formulation and simulation implementation. The potential value of the process-based modelling is that it provides a way of explaining system dynamics in terms of the mechanisms that drive the system.

Agent-based modelling is an appropriate modeling approach. There are two reasons. First, there is the same set of physical laws and biological laws that control the behaviour of agents. Second, the effect of the physical and biological laws governing individual interactions is analysed from the resulting emergent behaviour of the group of agents. In next chapter I model the physical aspect of agent behaviour, including the representation of agent, calculation of agent interaction and motility.

### 2.2.4 Model repurposing process

As described in Chapter 1 the research direction changed part-way through model

development and so the model was repurposed from vascular network simulation to tumour growth simulation, the detail of which is discussed in Chapter 5. This repurposing is accomplished by applying the CoSMoS process. The CoSMoS process [http://www.cs.york.ac.uk/nature/cosmos/about.html] describes a principled approach to scientific modelling and simulation: it provides a structure for managing and documenting the iterative development of a simulation, and gives scientists and simulation developers tools to reason - with an appropriate balance of confidence and scepticism - about how their simulation's results relate to the domain under study (Ye Li *et al*. 2014).

There are six phases in the CoSMoS modelling process, as shown in Figure 2.2.4.1. The first phase is ‘research context’, in which the high-level motivations and hypotheses are noted. The information is considered in the content of the ‘domain’, the subject of scientific research. The next phase is the ‘domain model’, in which scientific aspects about domain are understood. The domain model is the abstraction of domain. The domain model is used to develop the ‘platform model’, in which the simulation is designed based on understanding of the domain. Note that in the platform model, the design is also directed by the research context. The next phase is ‘simulation platform’, in which the simulation is encoded following the design in platform model and research context. The final phase is ‘results model’, in which the behaviour of simulation is observed and the result of simulation is recorded and analysed, typically in comparison with real data.

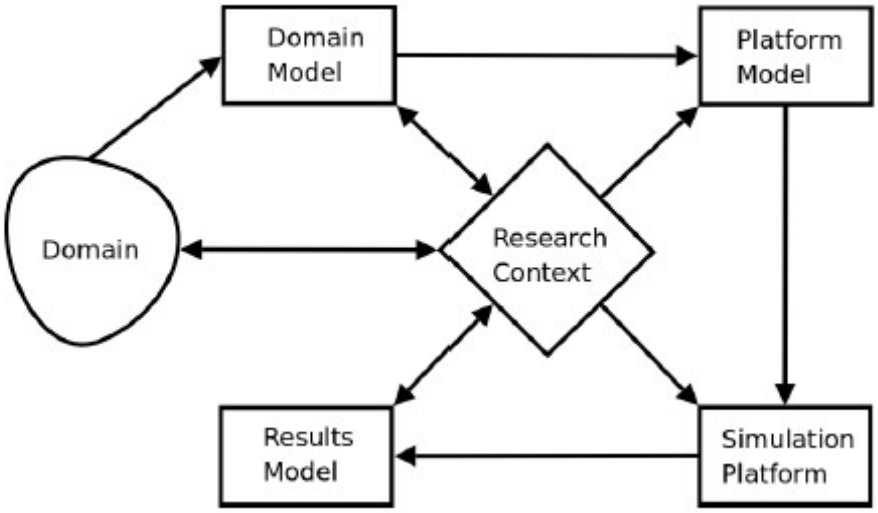



Figure 2.2.4.1 CoSMoS phases (Paul S. Andrews *et al*. 2010)

Here, CoSMoS is used to identify the changes that must be made to repurpose the model from vessel formation study to cancer cell interaction study. Note that neither of the models described here is built strictly under CoSMoS structure. Due to the complexity of both models, they are analysed following the CoSMoS structure phase after phase. By comparing the two models (vessel formation model and cancer cell interaction model) in each phase of CoSMoS structure, I understand the concept in models that are same and different. From the analysis, I also know the gap in my knowledge of the domain that must be filled by experimentation to calibrate the reused model. Drawing inspiration from the CoSMoS structure, I reuse the original model, including code structure, to the utmost extent, at the same time maintain confidence in the repurposed model (Ye Li *et al*. 2014).

## 2.2.5 Modelling approach as case study

In this section, three models are studied as exemplars of modelling approaches; each model focuses on one aspect of tumour morphology. All three models define rules for a single cell, and then examine the tissue-level behaviour revealed by group of homogeneous or heterogeneous cells. Following close inspection, these three models show how the simplifications and assumptions made in the model may affect the target model behaviour, as well as which aspects are crucial to each aspect of tumour morphology study.

In (Drasdo and Hohme, 2005), an agent-based model is built to explore how tumour growth is affected by glucose and cell-cell interaction in the early expansion stage. In this model, the cell cycle (see Section 2.4.1), cell proliferation and inter-cell force are modelled, in which the cell cycle can be affected by contact inhibition and the cell proliferation rate can be affected by concentration of glucose. The extracellular matrix (ECM) is not modelled.

In the model glucose is treated as a constant value outside tumour, and diffuses into tumour tissue with a fixed rate. The rate of consumption of glucose by cells is also fixed. The proliferation of cells is modelled as a response to a threshold of concentration of glucose $c_Q$, and the apoptosis of cells as a response to another threshold of glucose $c_{nec}$. In this way, the model shows dependency of cell growth to concentration of glucose $c$, in which $c$ is in the range of $c_{nec} \le c \le c_Q$. When $c$

exceeds the range, the growth of cell is independent of $c$.

The distribution of cell cycle time among cells in the population is initialised as a bell-shaped curve to reflect natural variation. As the cell growth is set to be inhibited by a threshold of cell-cell force, the growth of cells with neighbours is more likely to be slowed. Thus the real distribution of cell cycle time in simulation is deformed and has higher value with longer cycle time, as Figure 2.2.5.1 shows. Thus in macroscopic view, in the model, the tumour cell growth can be delayed by stress of other cells. However, at the edge of the tumour, the modelled cells are affected by less force and perform more divisions.

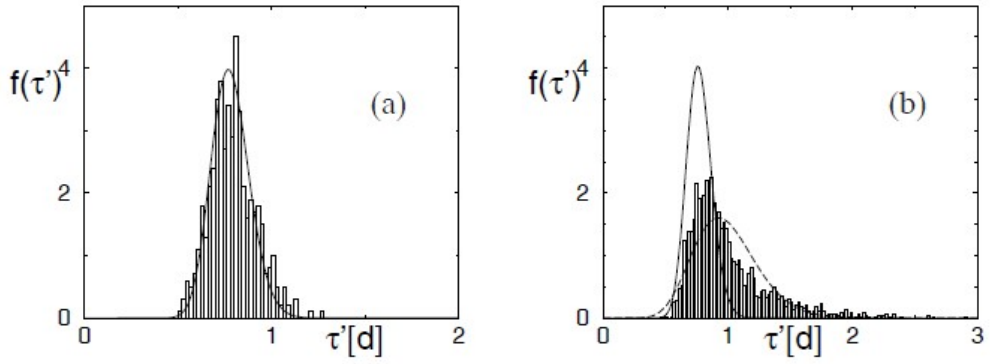



Figure 2.2.5.1 (a) distribution of cell cycle time in exponential growth period monolayers; (b) distribution of cell cycle time in linear growth period (dashed line). Comparing with exponential growth period (full line), the dashed line is deformed with more distribution on longer length of cell cycle. (Drasdo and Hohme, 2005)

This model is simulated and number of modelled cells is plotted as growth curve and compared with in vitro experiment data. In the experiments cells were planted to grow on petri-dish to form monolayer, or in suspension to form tumour spheroids. In monolayer the spatial distribution of tumour cells is two-dimensional, while in

spheroids the distribution is three-dimensional. Both simulations and experiments are carried out with various setting of concentration of glucose. This model is a good example of simulation of tumour growth with no vascular network (angiogenesis). Also the modelling explored how single cell behaviour (cell cycle length, proliferation) is linked to macroscopic morphology. It introduced a method to introduce effect of contact inhibition to cell growth, which produce a simulation result that is similar with in vitro experiment. This modelling approach also shows how the behaviour of a model is affected by the underlying assumptions and simplifications. Although this model demonstrates a complete modelling approach, the neglect of the effect of oxygen is suspect, as oxygen also plays an important part in tumour growth. The simplification of glucose (threshold of concentration) can be reused as oxygen except that the lack of oxygen (hypoxia) affects the cell cycle, an important part of this model. Indeed, the effect of hypoxia on tumour growth is studied in the following model.

My study aims to study the pattern formed by cancer cell, which means the cancer cells cannot be in a tight cluster, with which there is no structure to be studied. Thus not like this study, the contact inhibition to cell growth is not concerned in this thesis. However the cell cycle length and proliferation are important regardless of the pattern, and to study the cancer cell growth, these are the parts that to be considered. This study also shows that hypoxia can result a extended cell cycle, which should be considered in my hypoxia model. Please refer to Section 5.5.2 for more detail.

In (Shirinifard *et al*. 2009), the effect of angiogenesis on tumour growth rate, tumour size and shape is explored. In this model, angiogenesis is considered as a result of cell hypoxia. The model simulated the growth of tumour cells near simplified vessel structure, as well as angiogenesis. The vascular system is considered as a source of oxygen, so that simplifications are made to ignore the biological detail of vascular networks, for example, the smooth muscle of vascular walls and the blood flow inside. It is a lattice-based model including 3 types of tumour cell and ECM (modelled as a generalised cell with no cytoskeleton, no proliferation nor apoptotic capacity). The vascular system is modelled with 2 types of endothelial cells, one is the existing normal vascular cells, and the other is tumour induced vascular cells. The supply of oxygen and vascular growth factor VEGF-A are modelled as continuous concentration fields.

In this model, the normal tumour cell increases its volume and splits equally into two daughter cells when it doubles its original volume. When the normal tumour cell experiences hypoxia, it becomes a hypoxic tumour cells and secretes vascular growth factor. When the supply of oxygen is lower than a threshold the hypoxic tumour cells become necrotic tumour cells and lose volume constantly. The normal vascular cells in model have tight junction between each other. The tumour induced vascular cells, on the other hand, are initialised as inactive cells, which behave like normal vascular cells except that, when response to the vascular growth factor they turn to active

neovascular cells, which proliferate and tend to move in direction of high concentration of VEGF-A.

This model is not validated with experiments. It simulates the growth of tumour with and without angiogenesis, and the simulation result is compared with other models. Without angiogenesis, the simulation result shows a slowed cell cycle with the effect of hypoxia. In this model, the structure of existing vascularisation not only affects the growth rate of tumour, but also affects the shape of tumour. The tumour also shows interaction with the vascular structure, in which tumour pulls on and breaks the vascular structure, as Figure 2.2.5.2 shows. This model reinforces the importance of oxygen supply, and more importantly, it reveals that with proper simplification and assumption, the cell-level response has ability to show emergent behaviour. With validation with *in vitro* or *in vivo* experiments this model can be more valuable.

This model suggests that vessel formation and cancer cell growth are linked through hypoxia. The future work of my study, will be to simulate vessel formation together with cancer cell growth, although it is hard to operate *in vitro* experiments.

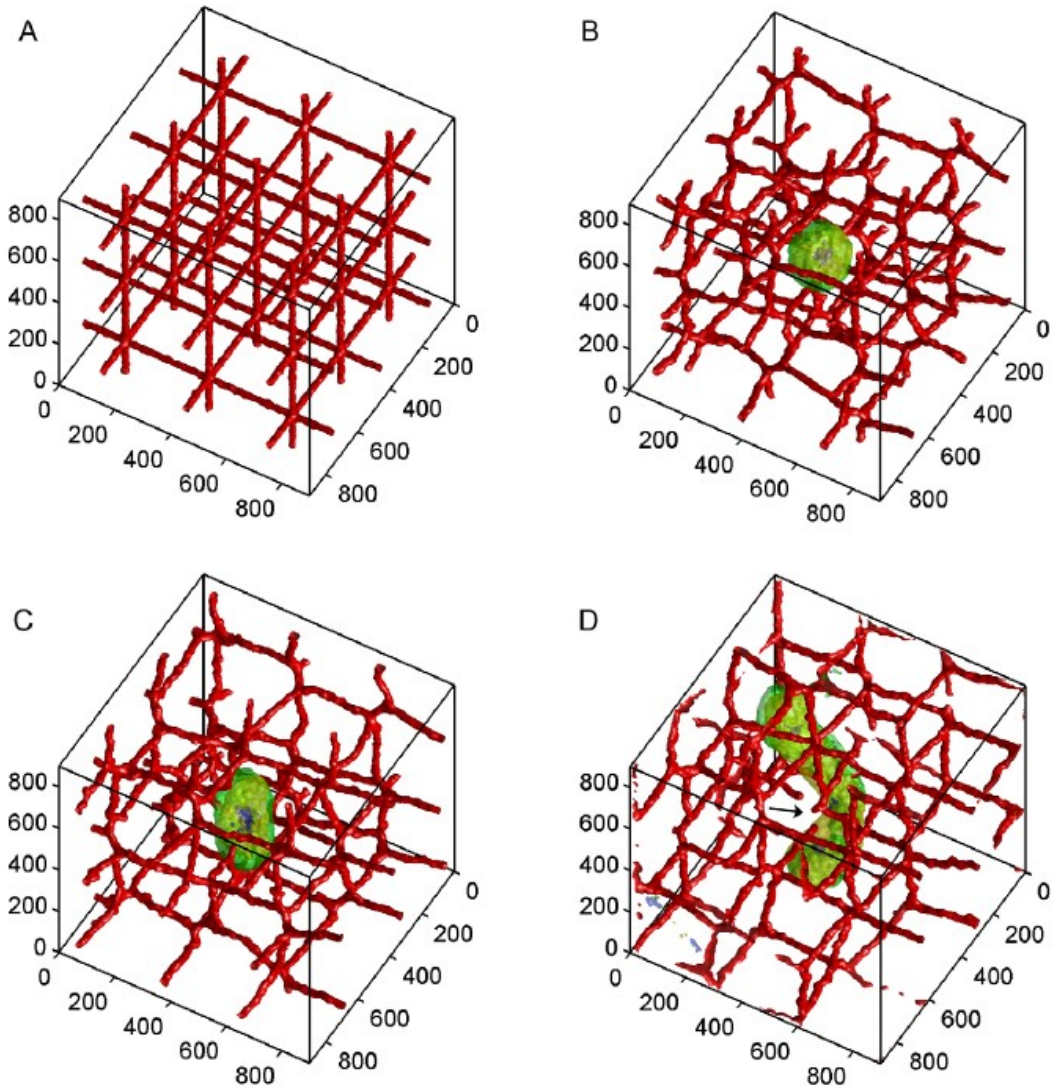



Figure 2.2.5.2 Shape of tumour in simulation near simplified vascular structure: (A) day 0; (B) day15; (C) day 30; (D) day 75. The tumour grows into a cylinder and finally breaks the vascular structure. (Shirinifard *et al*. 2009)

The invasion and migration of tumour is another important part of tumour study. In (Ramis-Conde *et al*. 2009), a model is built to study the influence of different protein pathways in the process in which cancer cells are transported into the vascular system. The pathways relative to cadherins (adhesion molecules) are modelled by ordinary differential equations. The physical interaction between pairs of cells and between a cell and extracellular matrix is modelled as friction forces, interaction force and random movement. The movement of tumour cell into vessel is modelled by the

Langeria equation. The modelled cell is considered to move at a low speed such that no acceleration is considered.

In this model the VE-cadherin and N-cadherin pathway is modelled. The VE-cadherin is considered to be the main factor that binds the endothelial cells together, while N-cadherin is the main factor that binds the cancer cell with an endothelial cell. The N-cadherin competes with VE-cadherin, thus the attachment of a cancer cell to an endothelial cell decreases the binding strength of this endothelial cell with other endothelial cells, as Figure 2.3.3 shows. In this way the cancer cell reduces the binding between endothelial cells and transport inside the vessel through the gap between unattached endothelial cells. The simulation of this model can reproduce the whole process of cancer cell attachment and transport.

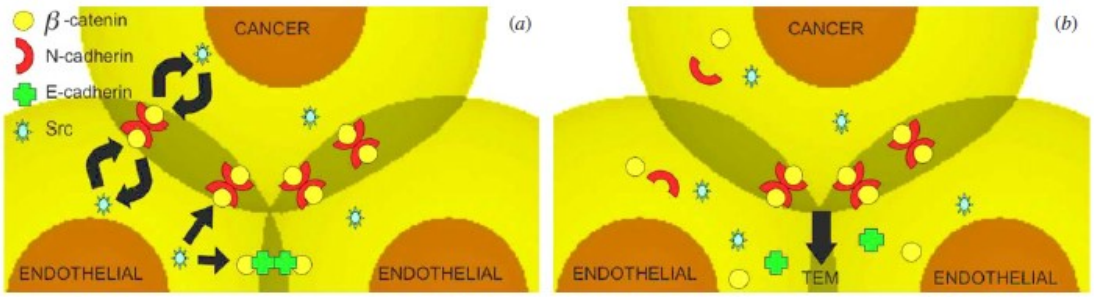



Figure 2.2.5.3 (a) a cancer cell make contact with pair of endothelial cells bound by VE-cadherin, and then the cancer cell is bound with endothelial cells by N-cadherin; (b) the N-cadherin in endothelial cells complete with VE-cadherin and result the loss of binds between a pair of endothelial cells. (Ramis-Conde *et al*. 2009)

The values of many parameters in this model are from other literature. To test the model, 4 genotypes are modelled to simulate effect of combination of pathways. The simulate result provides predictions about the length of time of transport of each

genotype.

The conclusion of this study is that the transport of cancer cells into the vascular system is mainly caused by physical properties (force, velocity) of the cell. Thus it enhanced the importance of cell-cell physical interaction to cell behaviour. It also shows another connection between physical interaction and cell behaviour (the other one is cell-cell stress and cell cycle). Therefore the physical interaction will be an important part of my study. Issues about physical model is discussed in Section 2.3.1, and the mathematical form of a physical model is built in detail in Chapter 3.

As a supplementary to previous 3 papers, Chen *et al*. 2009 provides another point of view on the oxidation in metabolism of cancer cells. This is not exactly a model, but a study to explain the reason of different response of cancer cell to hypoxia in *in vitro* and *in vivo* experiments. This study provides background of the character that oxygen plays in cell metabolism; also points out that same cell may have different behaviour in different environment.

## 2.3 Modelling issues

The above review, combined with additional literature, leads to the need to consider a range of issues in the modelling of endothelial cell interactions and cancer cell interactions. For endothelial cell, as it does not change shape or die during vessel formation, only physical issues are considered. For cancer cells, apart from physical

part, there are also biological aspects.

### 2.3.1 Physical issues

## Space: Lattice or Off Lattice

Lattice models are a type of models in which the position of modelled cells are constrained in grids or lattice. The modelled cells can move from one lattice to another one but cannot stay in-between. In models such as (Serini *et al*. 2003), each lattice can contain one modelled cell; in models such as (Andreas Deutsch, 2007) several modelled cells can stay in the same lattice. The advantage of lattice model is that it takes much less time to compute simulations than off-lattice models (Dirk Drasdo, 2007). However, the lattice spacing inherently provides a physical, artificial distance between cells. There is an assumption that cells do not overlap, although there is no justification for this. In addition, the direction and distance of cell motion are limited. The interactions between cells, and interactions between cells and environment, which is also lattice-based, is also limited (such as limit of number of cells that can have effect on each other at the same time, and the manner of inter-cell dynamic). These limitations in this exemplar are typical of those arising from lattice-based models.

There are also hybrid models comprising cells that move in a lattice structure and chemotaxins (the chemicals that can direct movement of cells) that is represented by a radiant field (Serini *et al*. 2003). In this way the interaction between cell and

environment (media, chemotaxins, etc.) is continuous and has an increased realism but at a computational cost.

In addition, the simplicity introduced by a lattice may affect the behaviour of model and produce artefactual phenomena (depending on the shape of lattice in particular model) in simulation. For example, in (page 9 of Andreas Deutsch, 2007 (page 37 of whole book)) simulations with square and hexagonal lattice produce significant different pattern: with the square lattice, the simulation result shows strong anisotropies, while the simulation result with hexagonal lattice shows isotropies.

Off-lattice models are also possible such as that described in (John C. Dallon, 2007). In this model the direction of cell movement is continuous. All the forces and torques act on centre of cell (agent). Unlike lattice model, there is no explicit assumption that cells do not overlap. Given an off-lattice model, there are two questions to answer: a) Are cells rigid or elastic? b) If cells are elastic, how to represent the degree of elasticity? These two questions are considered in Sections 2.1.4 and 2.1.6 respectively. Finally, off-lattice models can be computationally demanding and to address this problem performance may be improved by implementation optimization and parallel techniques, which is discussed in Section 2.5.

## Shape: Spheres and ellipses

Cell shape plays a significant role in cell behaviour (Rong Peng *et al*. 2011). Different types of cells are of various shapes and sizes. The shape and type may vary amongst the same type of cells or during the growth of a single cell. In order to model the shape of cells, certain simplifications are necessary. There are several levels of simplification. In the most simplified cases, cells are considered as identical particles in the models (Andreas Deutsch, 2007). In some models cells are considered as polymer of several elastic components (T. Newman, 2005), which is more detailed.

There are models that consider the cells as spheres such as (Jiujiang Zhu *et al*. 2006). There are several advantages: a) as the simplest 3-dimension shape, it is relatively simple to estimate spatial relationship between spheres; b) it is convenient to apply physical rules to a spheroidal shape. Despite these benefits, the disadvantage is also obvious: as a uniform shape, a sphere cannot represent anisotropic characteristics. The anisotropic interaction, however, may be necessary for morphological phenomena such as tumour invasion (R. Sodt, *et al*, 2014) and vascular formation (Hayashi *et al*. 2012). Cell polarity is mainly caused by the asymmetrical distribution of specific proteins on cell membrane. Cell polarity contributes to directional transport of cell and enable cell to sense surrounding cells and environment (Fernando Martin-Belmonte and Mirna Perez-Moreno, 2011). The polarity pathways are associated with tumour progression (Martin-Belmonte *et al*, 2011). In addition, polarity proteins are also linked with the Hippo pathway that regulates tissue growth

based on cell-cell contact pressure (Angela M. Liu *et al*. 2012) to limit organ size (B. Zhao *et al*. 2011).

In order to simulate anisotropic behaviour, the simplest shape is an ellipsoid. There are a couple of examples, such as (Eirikur Palsson, 2007), (Stolarska *et al*. 2009) and (K.J. Painter and T.Hillen, 2013).

## Inter-cell contact: Elasticity and adhesion

In Drasdo and Hohme, 2005 the physical interaction among tumour cells is considered as a factor affecting the distribution of the length of the cell cycle in a population of cells. The Physical interaction is also considered as driving mechanism of tissue level behaviour such as tumour invasion and intravasation, i.e. invasion of cancer cells into the lymphatic or vascular systems, in (Ramis-Conde *et al*. 2009). Thus the physical interaction among tumour cells plays important role on tissue morphology as an emerging behaviour. In this section, aspects of physical interaction are discussed.

Cell elasticity is important. When they are close enough and in contact with each other, real cells deform their shape. It is thus unrealistic to model cells as rigid particles. More importantly, in (Dirk Drasdo, 2007), it is argued that the cell elasticity is necessary to have force to affect motion of cells.

In some models, the cells are represented as deformable spheres with non-deformable

cubical cores (Dirk Drasdo, 2007). Although it is a good approximation for spheroidal cells, this is not possible for ellipsoid-based spheres. In order to keep a relatively simple representation of cell while assigning it degree of elasticity, another simplification is required. The cells in the model do not deform, but are permitted to overlap with each other to show degree of elasticity. The degree of elasticity is represented by the amount that two cells may overlap. In this way the representation of cell can be consistent.

As it is decided to use agent-based modelling approach, the model also keeps the potential of having each cell has different elasticity value.

Adhesive interaction is modelled in (Eirikur Palsson, 2007). It is considered not only one of causes of cell motion, but also a key mechanism to keep cells together to maintain normal or cancer tissue (Karin Sundfeldt, 2003). It is also one of the driving forces in a number of phenomena such as cancer cell invasion (Ramis-Conde *et al*. 2009). Therefore the adhesive interaction between cells is an important part for both cell motion and tissue morphology. As it is an inter-cell interaction that resulting tissue level phenomenon, it may become the bridge to link single cell behaviour with tissue morphology. Thus the adhesive interaction is an important part of models seeking to characterise cell-cell interaction.

To implement adhesive interaction a method is needed to estimate cell-cell interaction.

There are a group of methods called metropolis algorithms (Nicholas Metropolis *et al*. 1959). In these methods the spatial relationship between cells is considered to have a particular energy, and cells are considered to move to minimize the energy of the system (Nicholas Metropolis *et al*. 1959). The concept of minimum energy has been proved effective to explain the sorting of embryonic cells (M. Seinberg, 1963). With metropolis algorithm the energy of cell can be estimated; then the cell velocity is estimated from the energy.

There have been approaches using Boltzmann equations (Andreas Deutsch, 2007), Johnson-Kendall-Roberts (JKR) model (Dirk Drasdo, 2007), and there are similar approach using the Hertz-model (Hertz, 1882) for homogeneous elastic sphere interactions (Dirk Drasdo, 2005, 2007).

## Contact potential of agents

In section 2.4.3, I discussed the necessity of ellipsoidal shape of modelled cell. However when it comes to calculation of physical interaction, the ellipsoidal shape brings two problems. First, a concept is needed to describe spatial relationship between pair of modelled cells, especially how they overlap (if they do). Second, unlike spheroidal cells, a pair of ellipsoids of same position with different orientation may have different physical interaction; therefore it is inadequate to simply use the direct distance between pair of ellipsoidal cells as the term to represent their spatial

relationship.

To resolve these two problems, I introduce the concept of contact potential, which is a term to describe spatial relationship of two arbitrary ellipsoidal particles. This concept is also demanded by metropolis algorithms which are mentioned in introduction of Section 2.4.4 to estimate energy between pair of modelled cells. From the energy the force and other terms to describe the physical interaction are generated.

The concept of contact potential is used to calculate physical interaction of molecular in (Berne and Pechukas, 1972). In this study Berne and Pechukas introduced the Gaussian overlap potential (GOP) to represent a pair of elliptical particles. The GOP has a drawback, which is it can only be used for identical elliptical particles.

Then the study in John W. Perram *et al*. 1996 analyse the contact function and compare it with Gaussian overlap potentials (GOP) and proves the contact function shows correct extension of GOP. However it does not have the good computational features. (John. W. Perram *et al*. 1996)

In (Perram 1985) and (Perram *et al*. 1984) the closest distance between two arbitrary ellipsoids on given orientations is calculated and used to generate elliptic contact potential (ECP). Although ECP gives out correct result when two ellipsoids are in tangent contact (they just contact but do not overlap), it has several drawbacks. The

most important one is that ECP does not consider difference energy caused by different orientations (Paramonov and Yaliraki, 2005). In (Paramonov and Yaliraki, 2005), ECP is extended to be used for heterogeneous ellipsoids and it can be used for both rigid and elastic ellipsoids.

## Transfer potential to energy;

In introduction of Section 2.4.4, it has been clarified that the calculation of adhesive interaction is based on the assumption that the whole system tends to reach its minimum energy status. The contact potential is not energy in itself. Thus it needs to be transferred to energy, and then the energy is used to estimate physical quantities such as forces and velocity. According to John W. Perram *et al*. 1996, the GOP energy can be gained by replacing the square of the scaled length of spherical symmetric potential by the geometry relationship of two ellipsoids in Lennard-Jones potential (John Lennard-Jones, 1924).

## From contact/adhesion energy to force and torque;

To prevent cells from totally overlapping, the contact force is also added to balance the adhesion force. The contact force is zero if the pair of cells (agents) do not contact; then when they just connect, or in other words ‘are in tangent contact’, the contact force is still zero but increases as the pair of cells starts to overlap. There is an assumption here: the degree of cell overlapping represents the elasticity of cell. Therefore by tuning the scale of contact force and adhesion force in model, the

elasticity of modelled cell can be adjusted (to match experimental data if required).

From the need of contact force, it is also known that the contact potential is required in order to estimate value of contact force. Thus both contact force and adhesion force are estimated from the contact potential, though the specific equation may be different. In the remainder of thesis, the potential which is directly used to calculate contact force is called contact potential; and the potential which is directly used to calculate adhesion force is called adhesion potential and considered derived from contact potential with certain method.

According to the definition of force in classical physics, the force is (Goldstein *et al*. 2001). Thus the value of contact force equals the partial differential of contact energy against the displacement. Similarly the adhesion force equals the partial differential of adhesion energy. The torque equals the partial differential of energy against the direction of movement. There is another assumption here: that the density distribution within a cell is even and therefore in my model (see later) the forces are affected on geometry centre of ellipsoidal cell (agent).

## Fluid dynamics

There should be a reference states that cells can be considered as particles moving in fluid. I assume that each cell is considered under effect of resistance force which occurs when cell moves within the fluid (Beysens *et al*. 2000). When cells move in

the fluid, their maximum speed is limited (J. P. Rieu, *et al*. 2000; G. M. Walker *et al*. 2005). Note that although I do not consider acceleration, the velocity at each time step (in simulation) may vary. The forces are always balanced, so that the velocity at each time step in simulation is constant. Thus the Stokes resistance equations (H. Brenner, 1964) are applicable.

According to (Derjaguin *et al*. 1975; Brenner *et al*. 1980), for a symmetric ellipsoidal particle, the resistance force is in direct proportion to the viscosity coefficient. This process is straightforward but has disadvantages, the Stokes resistance formula is built for rigid particles, while in my model the cells are considered elastic, which may bring unknown errors.

## Environment: Gradients and Chemotaxis

Chemotaxis, i.e. cell movement in response to some chemical gradient, can occur in cells in response to internal (to the body) sources such as growth factor or from an external source, such as drugs. In response to gradient of chemotaxins, cells may change motion status (Serini, *et al*. 2003), or change growth status (Chen *et al*. 2009). In models such as (Serini, *et al*. 2003), the concentration of chemotaxis-inducing gradients is described by a radiant field.

## 2.3.2 Biological issues

To simulate behaviour of cancer cell, some of the biology properties, including growth, proliferation and apoptosis, of real cells need to be added together with physical part of model. In this section these parts are discussed.

## The cell cycle

The cell cycle is a sequence of events that happens in a cell leading to its division and replication (Geoffrey M Cooper, 2013. The cell: a molecular approach. The sixth edition). The cycle of a eukaryote cell contains three periods: interphase, mitotic phase and quiescent (Geoffrey M Cooper, 2013. The cell: a molecular approach. The sixth edition), as Figure 2.4.1.1 shows.

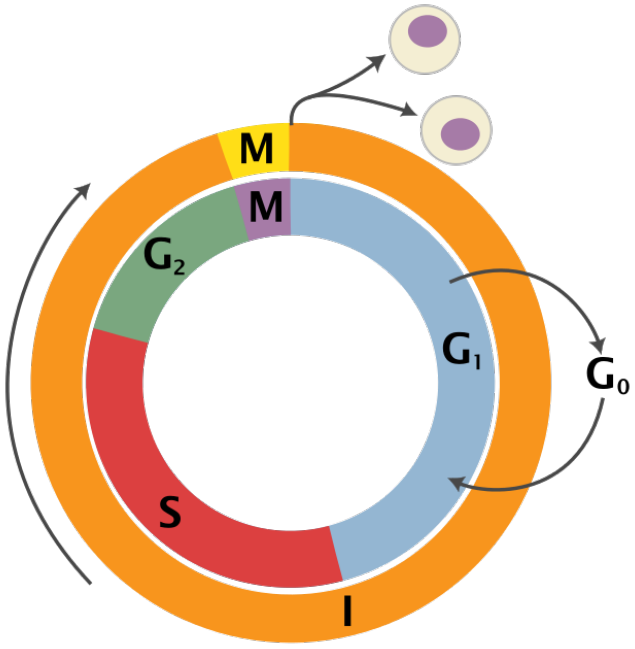



Figure 2.3.2.1 cell cycle phases [http://en.wikipedia.org/wiki/Cell_cycle]

The interphase can be divided into three stages: Gap 1, in which the cell increases in size, where the Gap 1 checkpoint (G1/S checkpoint) ensures that everything is ready for DNA synthesis; Synthesis stage (S stage), in which DNA replication happens; and Gap 2, in which the cell continues to grow, where the G2 checkpoint (G2/M checkpoint) ensures that the cell is ready to enter mitosis phase.

In animal cells, in the mitotic phase (or M phase) the cell stops growing, the nuclear envelope breaks down and the chromosomes in the nucleus separate into two identical sets. Then fibers form in the cell and pull sister chromatids to opposite polar ends of the cell and form two nuclei. After that, the nuclei, cytoplasm, organelles and cell membrane are divided into two cells that contain roughly equal shares of these components. The original cell is called the mother cell; the two cells are also called daughter cells and are typically identical to each other and to the mother cell as well. The Metaphase checkpoint ensures that the cell is ready to finish division.

Quiescent phase is also called Gap 0 stage, with respect of Gap 1 and Gap 2. Note that it is not a phase in which certain cell growth or division events happens, but a resting phase where cell stops growing. The cell may go into Gap 0 from Gap 1 stage and rest for long periods of time. Under proper condition, cell may transfer back to Gap 1 from Gap 0. The transition from and to Gap 0 may happen several times.

The apoptosis (programmed cell death) was firstly described in (Kerr *et al*. 1972). It is

a sequence of events leading to cell change and death. It is a quick progress, normally in which the cell shrinks, the cytoplasm increases density, condensation of chromatin, nucleus breaks, ended with cell breaking down. Apoptosis can be triggered by cell injury or other apoptotic signals. During this process the cells experience a series of complex biological and chemical reactions (Bicknell *et al*. 1994). The apoptotic pathways receiving the signals are shown in Figure 2.4.1.2.

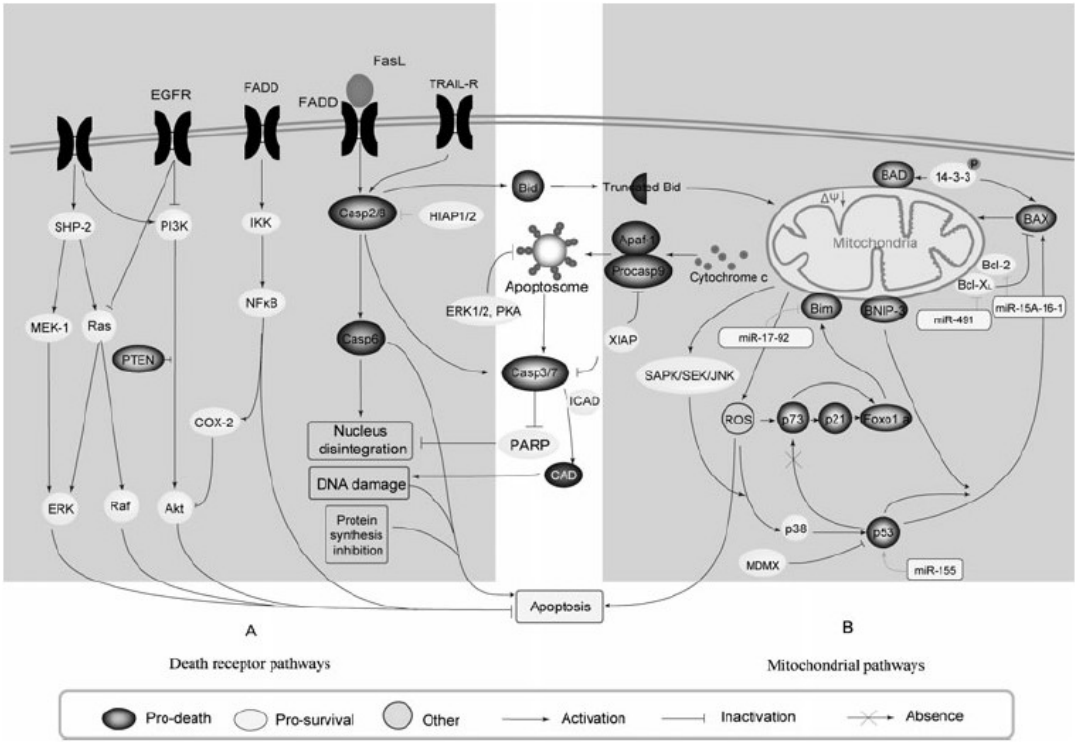



Figure 2.3.2.2 Two main types of pathways to induce apoptotic: death receptor pathways and mitochondrial pathways (L. Ouyang *et al*. 2012). The death receptor pathways are triggered by receptors on membrane when death stimuli occur. The mitochondrial pathways are controlled by intracellular signals.

For clarity in subsequent reading, the model developed here focuses on the physical parameters of single cell and how they affect the emergent behaviour of group of cells. Thus the size and shape of cell is the most important property to the model during the

cell growth process. In the physics-based model outlined in this thesis, the cell cycle is simplified. The change in shape and size of cells will be used instead of the real biological behaviour in the cycle of modelled cells. Thus the underlying biology is modelled phenomenologically.

The quiescent phase is simplified as an idle stage in the model, in which the size of the modelled cell does not increase. The interphase is simplified as the growing stage, in which means the size of the modelled cell increases. Finally, the mitotic phase is simplified as the split stage in the model, in which one modelled cell divides to two identical daughter cells. Apoptosis is simplified as a stage in model as well, in which the volume of cell is considered as zero. In this way the modelled cell may transfer between simplified stages. The detail of this simplified cell cycle is in Section 5.1. Also there is an assumption here: when the modelled cell is in the growth status, the increase of its volume is linear.

## Population growth

The rate that a group of cells increases their total number also represents the effect of environment on the cell. The total number of cells is an important system-level characteristic often considered in models.

A population growth curve can directly demonstrate the growth states of cells, and is

widely used in the researches of cell growth and regulation. In (Zhao *et al*. 2007) it is used to show the effect of YAP. It is used in (Deason-Towne *et al*. 2011) to demonstrate how MagT1 restores the growth of TRPM7−/− cells. Under perfect condition (sufficient nutrition, oxygen and space) the cell exponential growth in vitro experiment. Drawing number of cells overtime gives the exponential curve based on *e* (natural logarithm).

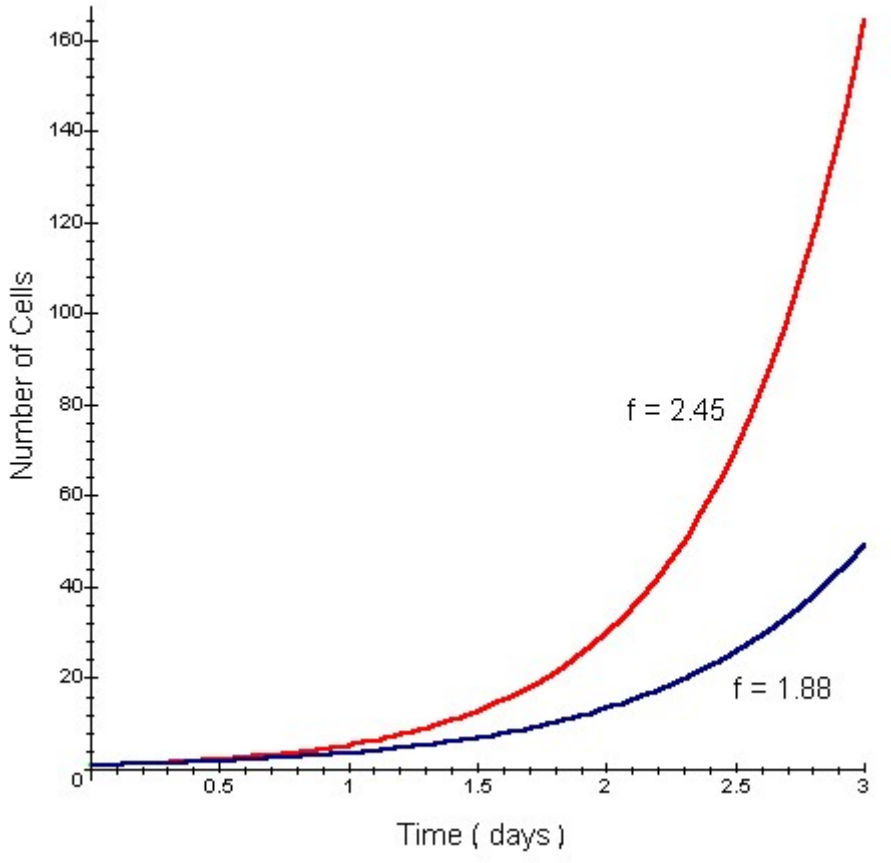



Figure 2.3.2.3 Exponential growth of cancer cell in 3 days . The value of f is frequency of cell cyces per unit time (1/day).
(http://www.tiem.utk.edu/~gross/bioed/webmodules/celldivision.html)

However in real *in vitro* or *in vivo* growth, cell growth does not exactly follow the exponential rule because they reflect events from a heterogeneous population of cells. The cell cycle of a single cell is regulated by various factors, and different cells

experience different factors and respond to those factors in different ways. For instance, on cell-cell contact, the protein kinase Hop (MST1/2) inhibits DNA replication (Zeng and Hong 2006). The cell cycle is arrested by gene p53 in Gap 1 on DNA damage (Tongyuan Li *et al*. 2012). A failed attempt of DNA restoration leads to programmed cell death (apoptosis). For a group of cells, the cell population is affected by both cell proliferation and apoptosis. If a group of cells are in a damaging environment (radiation, hypoxia, etc.), the effect of the environmental factors on each single cell is also demonstrated as the changes to the group population, resulting various shape of growth curves.

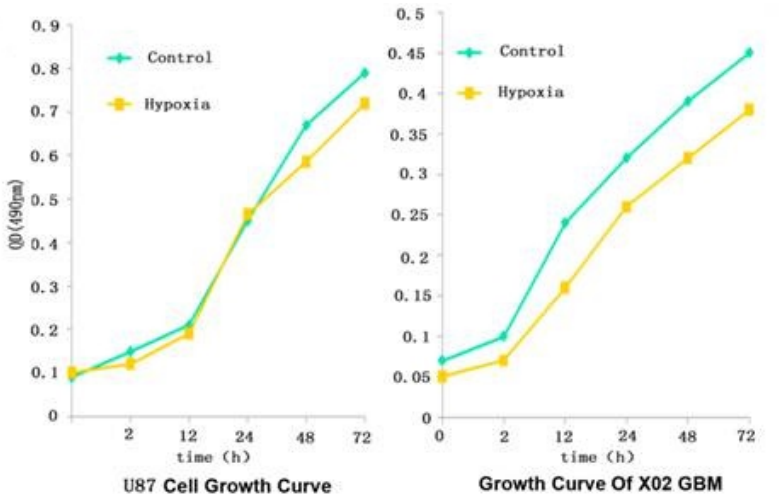



Figure 2.3.2.4 Growth curves of U87 cell and X02GBM cells. (Li, P., *et al* 2013) The shape of control curve (blue) and hypoxic curve (yellow) are different for both cell lines.

## 2.4 Software development approach

In Sections 2.2 the modelling approach and examples were discussed. In Section 2.3 the biological and physical aspects of single cell behaviour were outlined. An

agent-based model easily leads to object-orient programming (OOP) (https://en.wikipedia.org/wiki/Object-oriented_programming), in which one type of agent is represented by a ‘class’ that containing all data properties and functions of agent , while all the agents of the same class are ‘objects’ . Each agent object maintains its own properties by calling its functions, and in the in silico experiment a list of agent object is maintained.

In my software development approach, the physical properties of agent, such as position, velocity and force etc. Are packed into class ‘Ellipsoid’. Thus the endothelial agents are of class ‘Ellipsoid’. The program starts from the first agent A, and loops through all the Ellipsoid agents to calculate the total cell-cell force and cell-cell torque on it. Then the force and torque between A and substrate are calculated and added with cell-cell force and torque to form sum force and sum torque. After that the sum force and sum torque are used to calculate the velocity and angular velocity of agent. With the initial position and direction of agent A (considered as known), and calculated velocity and angular velocity, the new position and direction of agent A is then calculated and updated. The program then turns to the next agent, and then the next agent, until all the agents are updated, as shown in Figure 2.4.1.

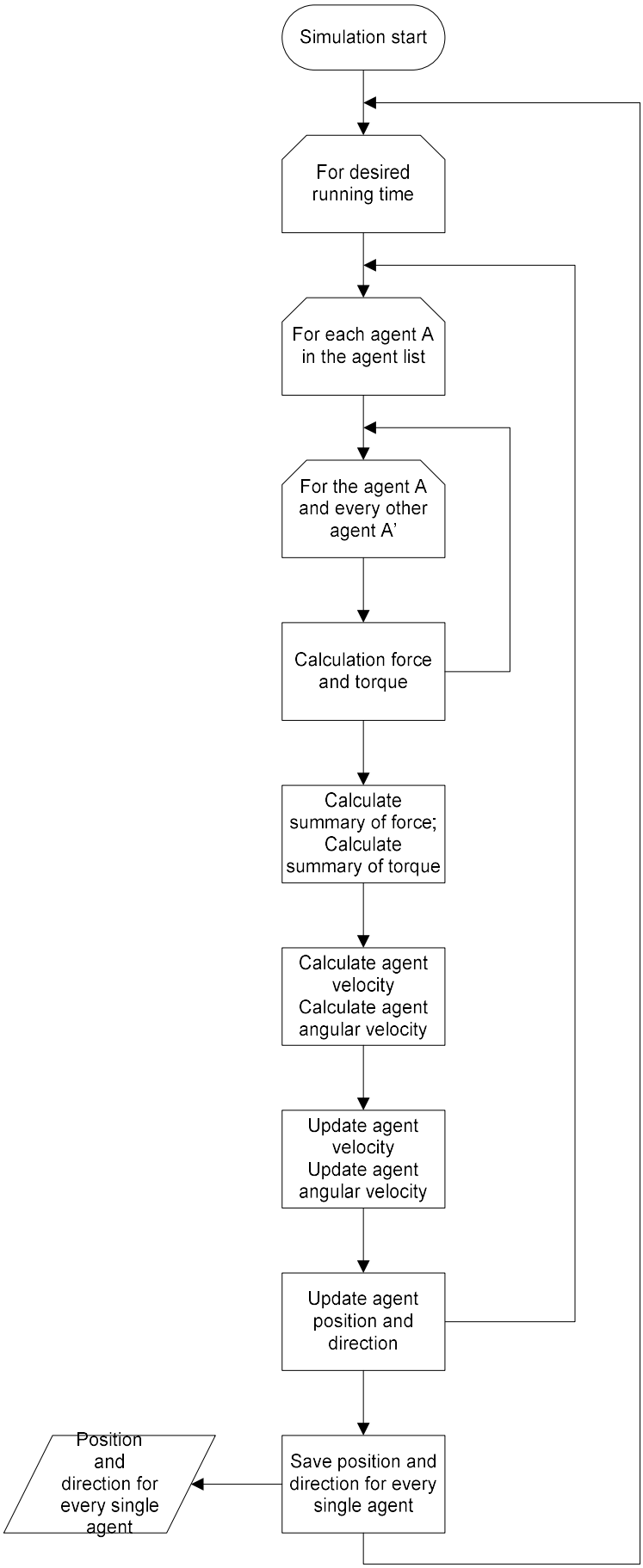

Simulation start
For desired running time
For each agent A in the agent list
For the agent A and every other agent A'
Calculation force and torque
Calculate summary of force; Calculate summary of torque
Calculate agent velocity
Calculate agent angular velocity
Update agent velocity
Update agent angular velocity
Update agent position and direction
Save position and direction for every single agent
Position and direction for every single agent

Figure 2.4.1 Process to calculate and update physical properties of all agents

The cancer cells, on the other hand, because they still have all the physical properties, are of a class named ‘Cell’ which is ‘derived’ from class ‘Ellipsoid’, and contains all the properties and functions of class ‘Ellipsoid’. The biological properties and functions are added to class ‘Cell’, so that the cancer cell can be aware of its age and growth status. Please refer to Appendix C: code and simulation issue for more detail of code structure.

When considering construction of agent-based models often computational limitations come into play. A key problem is that in order to calculate the physical forces impacting each agent, it is necessary to calculate those forces involved in the interactions between this agent and all other agents. Thus as number of agent increases, the increase in the number of calculations is not linear but squared. To reach realistic numbers in a population of agents in one simulation the running time of that simulation becomes unacceptable long.

Parallel computing technique is one potential tool to resolve this problem. By implementing these calculations concurrently, the simulation time can be significantly reduced, as Figure 2.5.1 shows. Note that the update of each agent state, including spatial position and any internal state, should happen after the all the calculation are done, therefore the calculation of physical properties of one agent does not affect its

position. In other words, calculation of properties of one agent should not affect calculation of properties of other agents so there are no order effects.

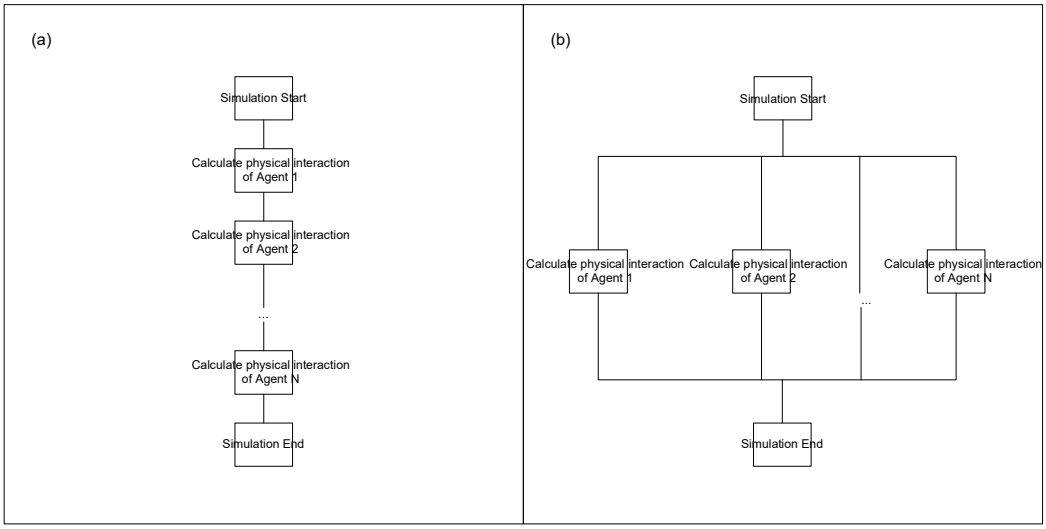



Figure 2.4.2 Calculation of physical properties with (a) and without data parallel (b).

Thread Building Blocks (TBB, https://www.threadingbuildingblocks.org/) is a parallel library released in 2006 by Intel Co. The latest version is 4.2. Unlike some other parallel interface such as MPI (http://www.mpi-forum.org/), it does not provide direct control over threads of execution (J. Reinders, 2007). Instead, TBB maps the calculation task onto threads automatically. It is useful in simulation in which a large amount of data needs to be processed in the same manner. Note TBB uses C++ code, which provides good links into the simulation code developed here.

Most importantly, the TBB library maps logical workload onto threads, which is a straightforward process for an agent-based model as the work load for each agent is already designed. At the same time, introducing TBB parallel library requires minor

change of original simulation code structure. For my agent-based model, the TBB library maps calculation of each single agent to multi-core processors, as Figure 2.5.2 shows.

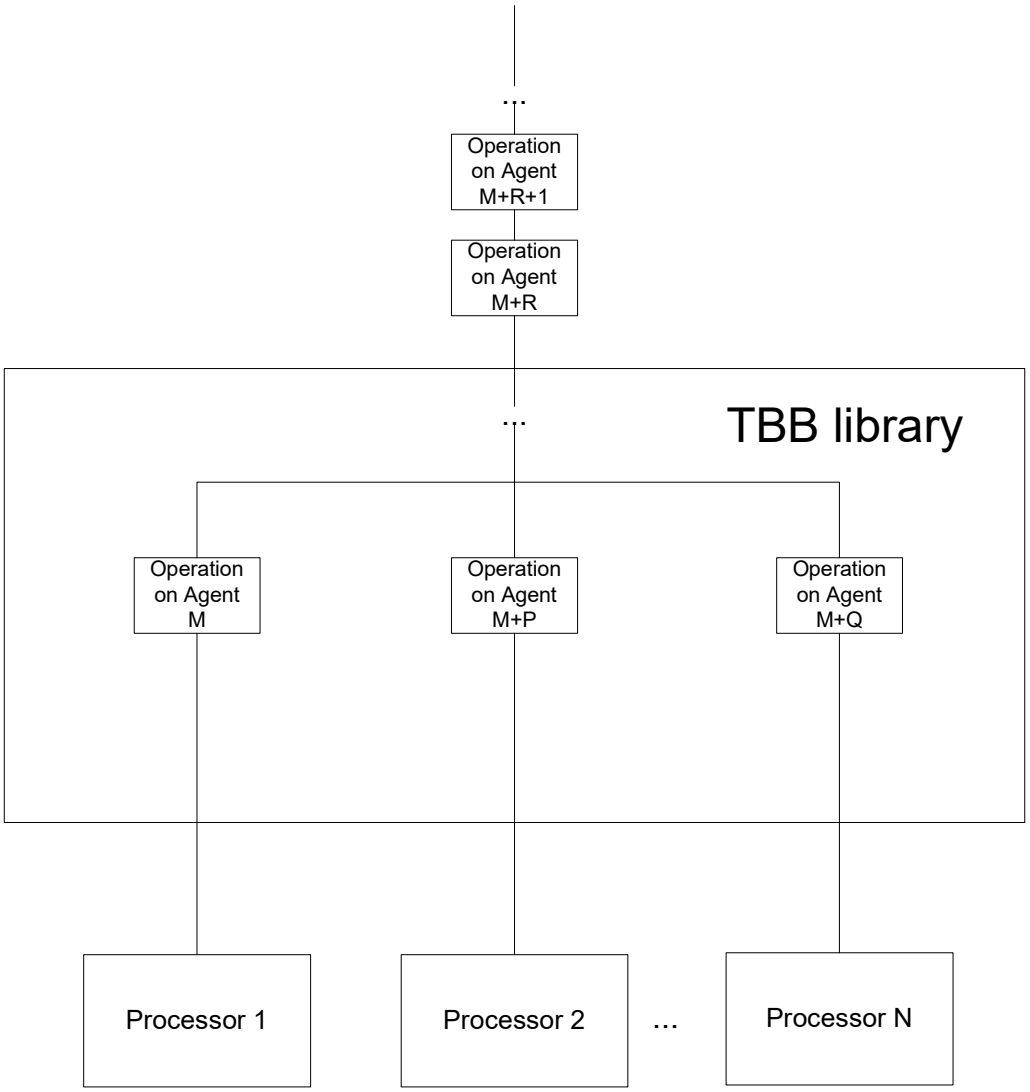



Figure 2.4.3 Concurrent operation of agent mapped onto N processors by TBB library

There is one potential problem with this solution: the TBB library is supported by Intel multi-core processors, on non-Intel processors the performance of simulation may not be as good.

## 2.5 Methods calibration methods

The spatial structure produced in simulation should be compared to the experimental data. In the early stage, the image of simulation can be directly compared with experimental data as a quality analysis, which helps to find mistakes in models. However it is far from enough to describe cell behaviour, so that is not efficient in comparison the simulation result of model with in vitro experiments, while the comparison is parts of process to fit experimental data into model (Merks *et al*. 2009).

In this section, quantity analysis tools are discussed, which will be used to explore the feature of the spatial structure, of both experimental data and simulation result. Also these tools will be used to calibrate simulation to experimental data. The calibration is discussed in Chapter 6.

### 2.5.1 Percolative transition

Percolation theory (Stauffer and Aharony, 1994) is used to study the connectivity of random generated structure, in which there is a control parameter $n_c$. For all $n < n_c$, the whole structure contains only un-connected clusters. As value of *n* increases, the structure is still un-connected, until $n = n_c$. When $n = n_c$, a transition happens and the structure is connected.

It is necessary to study percolative transition because (a) vessel network should allow blood to travle through to carry mutrients and oxygen to tissues, thus percolation is an

essential feature of vessel network; (b) the value critical exponents can be used as a feature of the system (Cavalli *et al,* 2007).

Percolative phase transition is used in Chauhan *et al*, 2012, in which the control parameter is the number density of cells. With low number density, the endothelial cells can only form isolated clusters, while with number density higher than $3000 cells / mm^3$ a connected structure is formed, as shown in Figure 2.5.1.

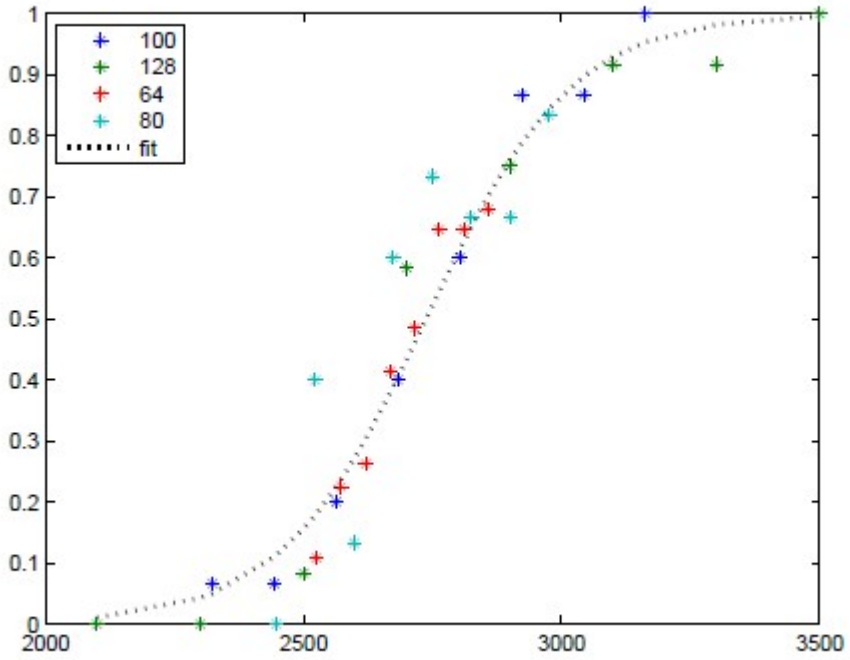



Figure 2.5.1 Percolation probability at various number density of cells (Chauhan *et al*, 2012). X-axis is number of cell per $mm^3$, y-axis is probability of connection. The percolative transition happens at number of density = 3000. '100', '128', '80' and '64' are size of observation window in four *in silico* experiments. Because all the *in silico* experiments have same dynamic feature, all the plots can be fitted to one percolation

curve.

## 2.5.2 Spatial statistic model

The history of point process can be tracked back to 1662 (Daley *et al*. 2007). Point process statistics is a tool to analysis the geometry structure formed by objects that distributed in one-, two-, or three-dimension space (Illian *et al*. 2008). In point process studies, the particles are abstracted as points and marks. The points are at location of objects, while the marks contain additional information about objects, such as shape and size.

## First order property

The intensity at a point in space can be considered as the mean number of events (points) in an area, i.e. a density measure. As equation (1) shows, $E(Y(d_s))$ is expected number of points, $d_s$ is area.

$$\lambda(s) = \lim_{ds \to 0} \{ \frac{E(Y(ds))}{ds} \} \qquad (1)$$

The intensity is a first-order property, which measures the distribution of event (agents in this study) (Bivand *et al*. 2008). It is a 'local' measurement, because it cannot give information about interactions among points. The interaction among points, for example, clustered or regular, is measured by second-order properties (Bivand *et al*. 2008).

Note that the equation in (1) is the theorical definition of the intensity, the value of intensity has to be estimated from experimental data or simulation result.

## Second order properties

Ripley's K-function and pairwise function are second-order properties of the spatial structure of the points. Second-order properties measure the strength and type of the interactions between points process (Bivand *et al*. 2008). There are higher-order properties, which are more complex and hence more difficult to estimate (Illian *et al*. 2008). Therefore higher-order properties are not discussed in this thesis.

The K-function is a method to measure a distribution based on distance. For a given point, the value of the K-function equals the number of points that can be found within a circle with this point as centre and radius *d*. If the intensity at this point is known, then the K-function should be equal to intensity multiplied by area of this point.

For a point process, the K-function at distance *d* equals sum of the K-function of every single point in this point process at distance *d* multiplied by intensity then divided by area of circle with radius of *d*.

$$K(d) = \int_{\rho=0}^{d} g(\rho) 2\pi\rho d\rho \qquad (1)$$

Note that the value of K-function has to be estimated from real value (gathered from

experiments or simulation).

For a distribution that is complete random or stationary Poisson process, the expected number of events exponentially increases over distance. Thus the K-function curve of a stationary Poisson process is an exponential curve and this serves as a baseline for interpreting plots. For an arbitrary curve, the part above an exponential curve represents the clustered distribution at the corresponding distance; on the contrary, the part of curve below the exponential curve represents the dispersive distribution at the corresponding distance.

The pairwise correlation function *g* is another tool to analyse a distribution. It relates to the K-function as follows

$$g(d) = \frac{K'(d)}{2\pi d} \qquad (2)$$

It is the probability of finding a pair of points at distance of *d*, divided by the corresponding probability for a completely random distribution (Poisson process) (Baddeley 2008).

The pairwise correlation function has been recently used in studies about the spatial structure of cell aggression (Binder *et al*. 2013. Treloar *et al*. 2014. Agnew *et al*. 2014), by analysing data from both experiments and simulation. These studies have illustrated the potential of pairwise correlation function to be used to explore the

feature of pattern of cells.

Same with the intensity and K-function, the value of pairwise correlation has to be estimated from experimental data or simulation result. The value of pairwise correlation $g(d)$ is always larger than 0. The points are considered clustered on distance $d$ where $g(d)>1$ and dispersed on distance $d$ where $g(d)<1$. For distance where $g(d)=1$, the point system is randomly distributed at distance $d$.

The pairwise correlation function can be drawn with distance $d$ as x-axis and value of g($d$) at $d$ as y-axis. The shape of pairwise correlation function can also provide information about the feature of the pattern. As shown in Figure 2.5.2, the solid line of $g(r)=1$ represents a typical completely random distributed point process. The dashed line ( $g(r)>1$ ) represents a typical clustered point process, because in a clustered system all the points are within a short distance, so that it is more likely to find a pair of points with short distance than points with long distance. The dotted line, on the other hand, represents a typical regular distribution, as for regular distribution, pair of points tend to appear with certain distance, which makes the pole.

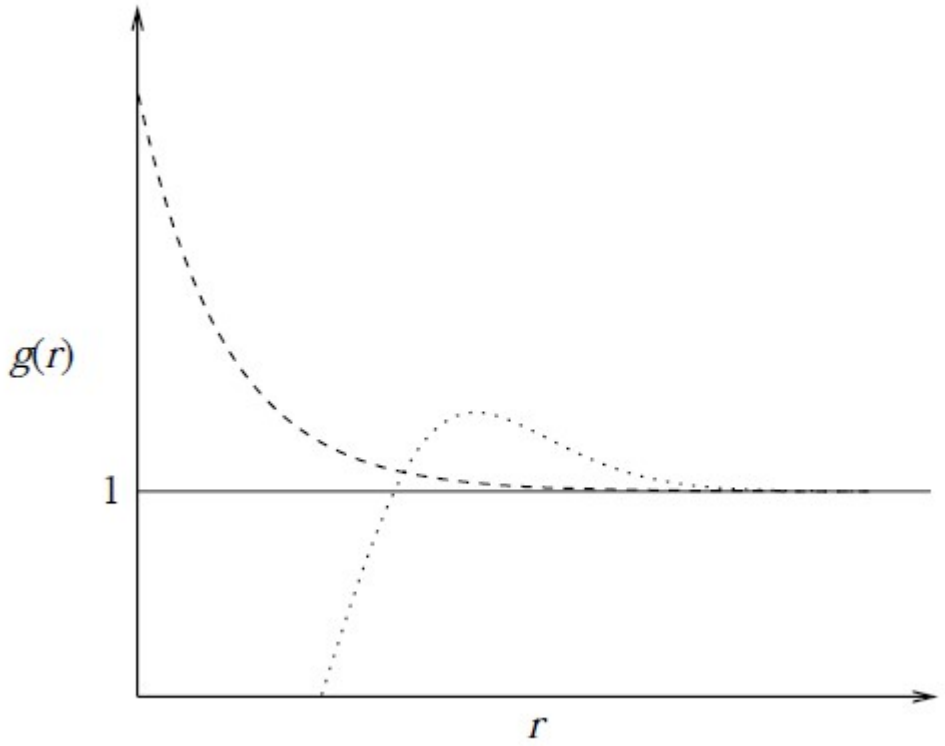



Figure 2.5.2 The typical shape of pairwise correlation functions of a completely random distribution (solid line), a clustered process (dashed line), and a regular process (dotted line) (Illian *et al*. 2008).

A typical point process with repulsion is shown in Figure 2.5.3, which shows multiple poles. The first pole shows the typical inter-point distance, while the second pole shows the typical distance between a point and its long-distance neighbours; the first non-zero minimum value that between the first and second poles suggests the typical distance between a point and area with small number of points (Illian *et al*. 2008).

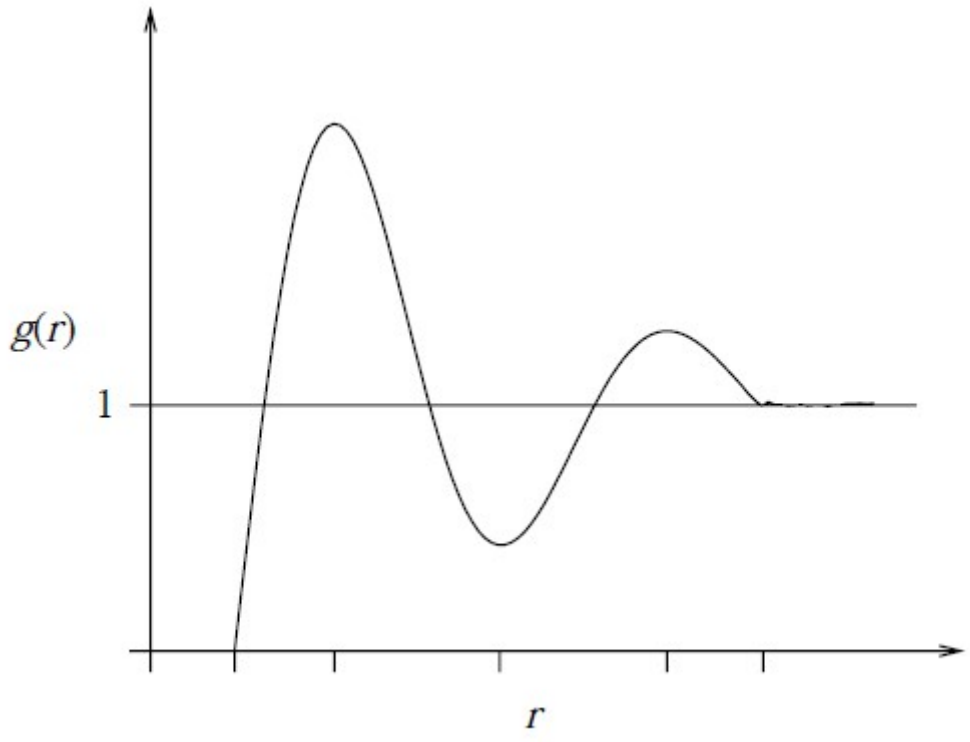



Figure 2.5.3 A typical pairwise correlation function of repulsive distribution (Illian *et al*. 2008).

## Conclusion

As discussed in Section 2.2, the physical aspect of agent behaviour, including the representation of agent, calculation of agent interaction and motility will be modelled following agent-based modelling approach in the next chapter. Then in chapters 4 and 5, the biological aspects specific to two different biological systems are discussed, including the set of biological assumptions and the experiment designed to calibrate the biological model. Based on the same experiment and the results shown in chapter 5, the calibration of the physical aspects, in addition to the biological aspects, is discussed in the chapter 6 using pairwise correlation function.

According to Section 2.3.1, this model is built based on the assumption that agents move until the whole agent system reaches its lowest energy status. In Chapter 3 I begin my account of the model with the calculation of the energy of system. The agents are represented as ellipsoidal particles, and the reasoning behind this is outlined in Section 2.3.1. Therefore in Section 3.1, I start with the mathematical representation of an ellipsoid, including its shape, size and direction. Then, taking an example of a two-agent system, a method is demonstrated in Section 3.2 to calculate the potential from the geometry information and relative position of the two ellipsoidal agents, which is an extension of derivations in (Perram *et al*. 1996) and (Paramonov and Yaliraki, 2005).

As one of the purposes of this model is to study the effect of cell shape to migrant pattern as emergent behaviour, the particle is too simple; the polymer of elastic components is too complex. The ellipsoid is chosen to represent geometry shape of cell. As ellipsoid is anisotropic shape, instead of distance between pair of cells, the contact potential is used to describe cell-cell interaction. Then the early simulation shows that the motion of modelled cell is non-realistic, which results the usage of Hertz formula. As Hertz model work with spheroidal particles, an extension is made to use Hertz formula for heterogeneous and for ellipsoidal particles. The detail of extension is discussed in Chapter 3.

The adhesive interaction is abstracted as a force which exists between each pair of

cells. The adhesive force affects both cells equally and tends to drag them to move towards each other. There is adhesion force between cell and petri-dish as well (Manoussaki *et al*. 1996). Similarly, adhesion force between one single cell and petri dish moves the cell towards surface of petri dish. I assume that it can be affected by both environmental factors and intra-cell factors.

I not only need to estimate the cell-cell interaction, the cell-petri-dish interaction is also required to be estimated. Because the interaction of two identical cells can be considered mirror symmetry, the interaction between one cell and the substrate plane, can be considered as the interaction between this cell and its mirror image. Thus I consider it is proper to calculate the interaction between cell and petri-dish in the same way of calculating pair of cells. There are detailed ellipsoid-plane interaction formulas (Brenner and Gajdos, 1980), but my method keeps the consistency of calculation of interaction of cells. At last, in the model I assume all the cell dynamic happens on surface of petri dish. Thus a monolayer of modelled cells is arranged on the same plane. This assumption requires additional rule applied to cells: I have to assume the interaction between modelled cell and plane is much larger than interaction between pair of two cells to prevent them from piling up (it also restricts the behaviour of the model).

The real cell increases its size during growth, and it divides to two daughter cells when its volume is almost doubled. To simplify this process, the modelled cells

cannot change their shape. But they can gradually increase their volume until it is exactly doubled. The scale of overall shape is maintained. In this way the modelled cells can be used to simulate growth behaviour and I can focus on the growth status of modelled cells. This part is discussed in Chapter 4.

In this model, there are effects of chemotaxins such as 5-FU, however I assume that it is evenly distributed in petri dish and has equal effect on every single cell (agent). Therefore instead of having an evenly distributed radiant field, the mechanism of drug is simplified and equally added to rules that define the behaviour of each single cell (agent). In real cell dynamic (or treatment) the chemotaxins cannot be extremely evenly distributed, the evenly distribution assumption is another simplification, which helps me to focus on the effect of certain concentration of drug to the cells.

Also in Section 2.2.5, the effect of oxygen on tumour cell metabolism is studied in (Shirinifard *et al*. 2009), which slows cell cycle process and relates to secreting of chemotaxins such as VEGF-A. Thus the effect of hypoxia should be included in the model. While (Y. Chen, *et al*. 2009) provides view of oxygen metabolism in biological point of view in Section 2.2.5, the method to model hypoxia is discussed in Chapter 5.

Based on discussions in Chapter 2, I make following assumptions:

| | Explicit assumptions: |
|---|---|

| | | |
|---|---|---|
| 1 | All the cell dynamic happens on surface of petri dish; | To produce a intuitive pattern |
| 2 | The interaction between modelled cell and plane is much larger than interaction between pair of two cells; | Scianna *et al*. 2013 |
| 3 | 5-FU is evenly distributed in petri dish and has equal effect on every single cell; | For simplicity of calculation |
| 4 | Ellipsoidal shape; | R. Sodt, *et al*, 2014. Hayashi *et al*. 2012 |
| 5 | The ratio of three semi-axis of ellipsoid keeps the same during the cell growth (the shape of cell does not change but the size changes); | For simplicity of calculation |
| 6 | Each cell is considered under effect of resistance force which occurs when cell moves within the fluid; | Beysens *et al*. 2000 |
| 7 | The degree of cell overlapping represents the elasticity of cell; | To keep a relatively simple representation of cell |
| 8 | The density distribution within cell is even and therefore in my model the forces are affected on geometry centre of ellipsoidal cell; | For geometric simplicity |
| 9 | The fluid in which cells grow and migrate is still, i.e. the fluid has no velocity or angular velocity. | It is the case in *in vitro* experiments |

| | | |
|---|---|---|
| 10 | The growth of cell volume is linear; | For simplicity of calculation |
| 11 | No upper limit of total number of cells; | The length of experimental time is short, in which the cells have not reached the upper limit of total number. |
| 12 | The cell age obeys normal distribution; | Drasdo and Hohme, 2005 |
| | Implicit assumptions: | |
| 13 | Only contact force, adhesion force and resistance force are significant forces on the cell (density of cell is simiar with water thus gravity does not need to be considered). | Grover *et al*. 2011 |
| 14 | Contact force can be computed using the Perram-Wertheim approach; | John W. Perram *et al*. 1996 |
| 15 | Cells move at very low speed so their acceleration approaches zero and the forces upon them are balanced (precondition of using Stokes resistance law); | J. P. Rieu, *et al*. 2000. G. M. Walker *et al*. 2005 |
| 16 | Resistance force can be computed using Stokes' law, and the known properties of the fluid medium. | H. Brenner, 1964 |

# Chapter 3 Inter-cell interaction model

## Introduction

Following the reviews in Section 2.3, the physical interaction plays an important role in tumour morphology. In Chapter 3, the detailed physical interaction is modelled. Then in Chapter 4 and 5, cell cycle and population growth of two different biological systems are discussed, including the set of biological assumptions and the experiment designed to calibrate the biological model. Based on the experiment result, the cell cycle and population growth are calibrated in Chapter 5, and then the calibration of the physical aspects is discussed in the Chapter 6.

As discussed in Section 2.4.1, this model is built following the agent-based process. Thus the physical interaction is presented in the form of a set of rules that act on each agent. In the other words, the set of physical rules are considered as the physical behaviour of each agent, while the result of interaction considered as emergent behaviour of group of agents. In Section 2.4.3, the necessity of considering real cells as ellipsoidal particles is explained. Thus in Section 3.1, the building of physical model starts with representing the geometry information, including its shape, size and direction, of an arbitrary ellipsoid. Although it is much easier to model all the agents as identical particles, the tumour tissue may contain various types of cells, which may in various phases of cell cycle. Due to the heterogeneous feature of tumour tissue, this

model is required to be able to represent mixed agents with various shape, size and cell cycle.

According to Section 2.4.4, this model is built based on the assumption that agents move until the whole agent system reaches its lowest energy status. Then in Section 3.2, taking example of a two-agent system, the mathematical form of potential between pair of arbitrary ellipsoidal particles is derived from the geometry information and relative position of the two agents, which is an extension of derivations in (Perram *et al*. 1996) and (Paramonov and Yaliraki 2005).

According to section 2.4.4, the energy can be derived from potential by using Metropolis algorithm; following this approach, and detailed in section 3.3 and 3.4, the potential is firstly transformed to contact energy and adhesion energy. According to its physical interpretation, the contact energy is transformed to contact force, and the adhesion energy is transformed to adhesion force. Similarly the contact torque and adhesion torque is calculated from the corresponding energy in Section 3.5.The adhesion force and torque work as contraction effectors, while the contact force and torque work as repulsion effectors. Similar with [Ramis-Conde, et al. 2009], I assume that the cells move in a low speed that there is no acceleration, thus the sum of force is directly used to calculate velocity of agent; and the sum of torque can be used to calculate angular velocity of agent, as explained in Section 3.6. The velocity and angular velocity respectively in 3.6, which at last are transformed to position and

direction of each agent, are shown in section 3.7.

As I can see, the calculation of physical interaction for each agent is relatively complex, which is another reason that parallel computing technique is necessary for large-scale simulation.

Finally, to ensure the correction of physical interpretation between all the transforms and estimate the value of constants in the formula, a dimension analysis is carried out. The correction of dimension of equations in the model is checked in Section 3.9, and the values of parameters in the model are estimated in Chapter 4 and 5 separately as the model is used in these two chapters for different purpose.

## 3.1 Representation of a cell

Given my rationale for representing cells as ellipsoids, it is necessary to represent the cell geometrically. As Figure 3.1.1 shows, an ellipsoid has one centre point, and three semi-axes. Note that the ellipsoid is axisymmetric along the three semi-axes: I assume that the mass is evenly distributed over the ellipsoid so all of the forces can be considered as acting on the centre of the ellipsoid.

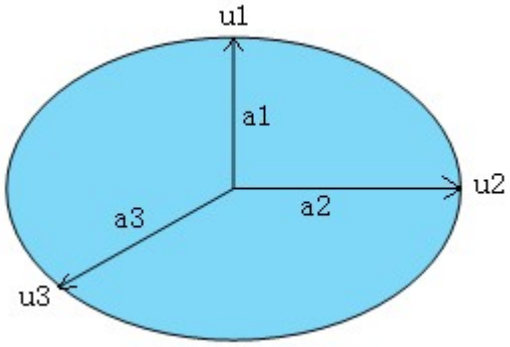



Figure 3.1.1 An ellipsoid. The length of the three semi-axes are $a_1$, $a_2$ and $a_3$ ($0 < a_1 < a_2 < a_3$), the direction of the three semi-axes are $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$.

I consider the three semi- axes as three vectors in space, so that they can be written in the form of three unit vectors to represent their direction, and three values to represent their length. The three unit vectors are refered as $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$, and their corresponding length are refered as $a_1$, $a_2$ and $a_3$. From the characteristics of an ellipsoid, it is known that $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ are orthometric. For simplicity, I label the shortest length $a_1$ and longest $a_3$, i.e. $0 < a_1 < a_2 < a_3$.

To represent the ellipsoid in a single expression for convenience, $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ are combined together as a matrix:

$$\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \quad (1)$$

In which each column represents the direction of one semi-axis. Further, $a_1$, $a_2$ and $a_3$ should be combined with their corresponding direction, so that the ellipsoid A can be written in the form of $\mathbf{A} = a_1^{-2}\mathbf{u}_1 \otimes \mathbf{u}_1 + a_2^{-2}\mathbf{u}_2 \otimes \mathbf{u}_2 + a_3^{-2}\mathbf{u}_3 \otimes \mathbf{u}_3$, in which $\otimes$

is the outer product. As per the matrix form used in (Paramonov and Yaliraki 2005)

$$\mathbf{A} = \mathbf{U}\bar{\mathbf{A}}\mathbf{U}^T = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \begin{pmatrix} \frac{1}{a_1^2} & 0 & 0 \\ 0 & \frac{1}{a_2^2} & 0 \\ 0 & 0 & \frac{1}{a_3^2} \end{pmatrix} \begin{pmatrix} u_{11} & u_{21} & u_{31} \\ u_{12} & u_{22} & u_{32} \\ u_{13} & u_{23} & u_{33} \end{pmatrix} \qquad (2)$$

In which $\mathbf{U}^T$ is the transposition matrix of the matrix $\mathbf{U}$, and $\bar{\mathbf{A}}$ is the eigen matrix of matrix $\mathbf{A}$. It is obvious that the matrix $\mathbf{A}$ is an orthometric matrix. Another observation is that the eigen matrix of an ellipsoid $\mathbf{A}$ with semi-axes $a_1$, $a_2$ and $a_3$ has the form of

$$\begin{pmatrix} \frac{1}{a_1^2} & 0 & 0 \\ 0 & \frac{1}{a_2^2} & 0 \\ 0 & 0 & \frac{1}{a_3^2} \end{pmatrix} \qquad (3)$$

As matrix U represents the direction and $\bar{\mathbf{A}}$ represents the length of the three semi-axes of ellipsoid, consider an ellipsoid with its centre at the origin, all its geometry information is included by matrix A. In other words, matrix **A** represents an ellipsoid at the origin with direction **U** and semi-axes length of $a_1$, $a_2$ and $a_3$. This form of ellipsoid is used in all the following sections.

A point of note is the need for $\mathbf{U}^T$. The matrix **U** contains the directional information. According to the cross product rule, $\mathbf{U}^T$ is added so that the determinantal expansion

of matrix form of the ellipsoid is the same as the expansion of the algebraic form of the ellipsoid.

## 3.2 Contact potential

For those models in which the geometry of cells is represented by a sphere, in order to tell the spatial relationship between any two cells, the distance between them is to be measured and compared with the sum of their radii. If the distance is larger than sum of their radii, then the two cells are not in contact; conversely, if the distance is less than this sum, the two cells are partly overlapped; and if the distance equals the sum, the two cells are just contact, or 'externally tangent', as Figure 3.2.1 shows.

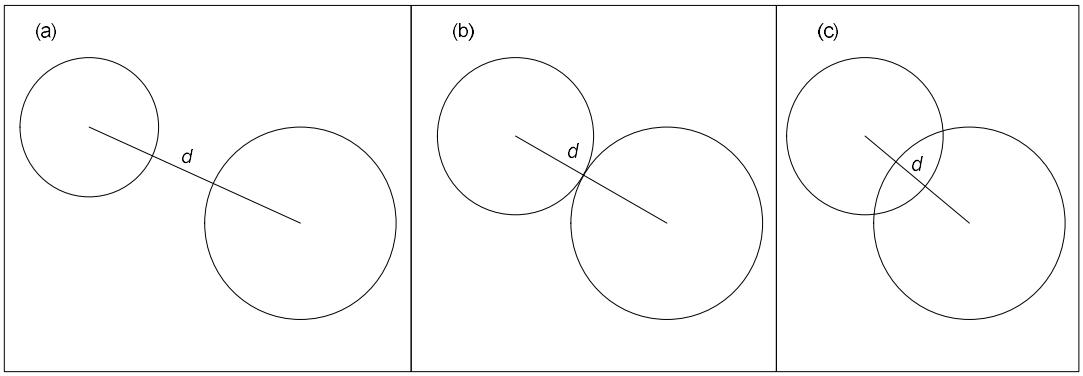



Figure 3.2.1 Two non-identical spheres. (a) For non-touching spheres, the distance between the centres is larger than the sum of radii of the spheres; (b) for the just-touching spheres, the distance between the centres equals the sum of radii; (c) for overlapping spheres, the distance between the centres is less than the sum of radii.

However with an ellipsoid model, this simple approach is no longer valid, as orientation plays an important role in spatial relationship. As figure 3.2.2 shows, while they are at fixed position, whether two ellipsoids are in contact or not depends on their orientation.

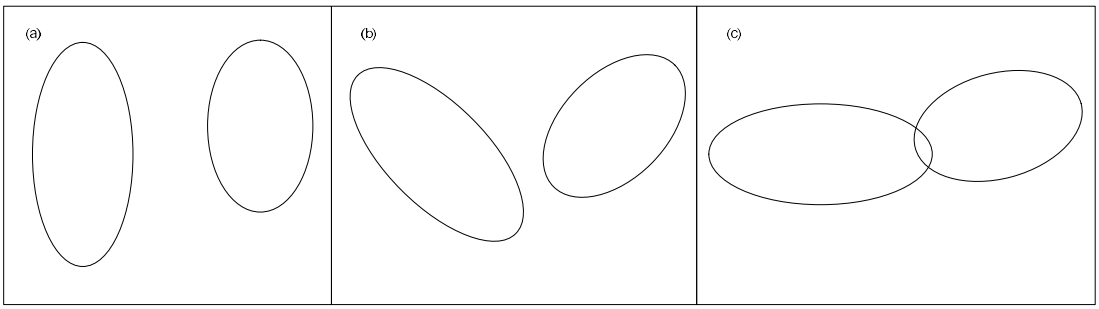


Figure 3.2.2 Two ellipsoids of arbitrary size at same position with various orientations. In (a-c) the distance between the centres of the pair ellipsoids is the same, but with various orientation the ellipsoids may be overlapped or not: (a) two ellipsoids lie in parallel, not in contact with each other; (b) two ellipsoids are not parallel and not in contact with each other; (c) two ellipsoids are not parallel and partly overlap.

In order to calculate the spatial relationship of two ellipsoids, the concept of contact potential is introduced, which is discussed in section 2.2.1. The contact potential is the most important term in the model, as all the forces and torques are calculated from it. It means a certain level of energy. As shown in Figure 3.2.3, the dotted lines are potential surfaces around the ellipsoid cell, and any point on the same surface has the same potential (note the continuous space is discretized here for explanatory purposes).

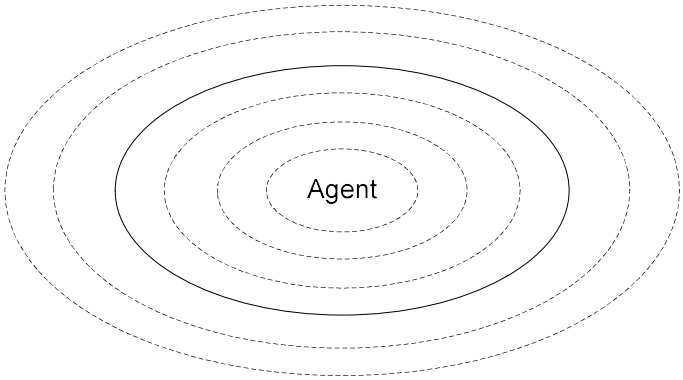


Figure 3.2.3 Potential surface of ellipsoidal agent. The solid line is the surface of the agent. The dot lines are potential surfaces.

With contact potential, the relationship between an arbitrary point in space and the ellipsoid can be described. First the value of potential of that point is calculated (as detailed below). The larger this value, the further away that point is from the ellipsoid. Inversely, the smaller this value is, the closer that point is to the ellipsoid. However potential is a relative value, which means that only the change of potential can be measured and valued. This relative nature of potential is not convenient to use for descriptions and so, for convenience here, let the centre of ellipsoid have potential value of zero, and the surface of the ellipsoid have potential value of one. Given this assumption, the contact potential is treated as follows.

Assume an ellipsoid **A** can be represented as $\mathbf{A} = a_1^{-2}\mathbf{u}_1 \otimes \mathbf{u}_1 + a_2^{-2}\mathbf{u}_2 \otimes \mathbf{u}_2 + a_3^{-2}\mathbf{u}_3 \otimes \mathbf{u}_3$. The contact potential of an arbitrary point **r** from $A$ can be represented as $F_A(\mathbf{r} - \mathbf{r}_A) = (\mathbf{r} - \mathbf{r}_A)^T \mathbf{A} (\mathbf{r} - \mathbf{r}_A)$ (J. W. Perram *et al.* 1985), where $\mathbf{r}_A$ is the position vector of the centre of ellipsoid **A**, and $\mathbf{r} - \mathbf{r}_A$ represents the vector starting from centre of ellipsoid **A** pointing to the arbitrary point **r**. The spatial relationship between **r** and **A** can be easily determined as:

$$F_A(\mathbf{r} - \mathbf{r}_A) \begin{cases} < 1 \ for \ \mathbf{r} \ \text{inside} \ A \\ = 1 \ for \ \mathbf{r} \ \text{on the surface of} \ A \\ > 1 \ for \ \mathbf{r} \ \text{outside} \ A \end{cases} \qquad (1)$$

In which $\mathbf{r}$ is the vector that represents the position of the arbitrary point, $\mathbf{r}_A$ is the vector that represents the position of centre of ellipsoid $A$.

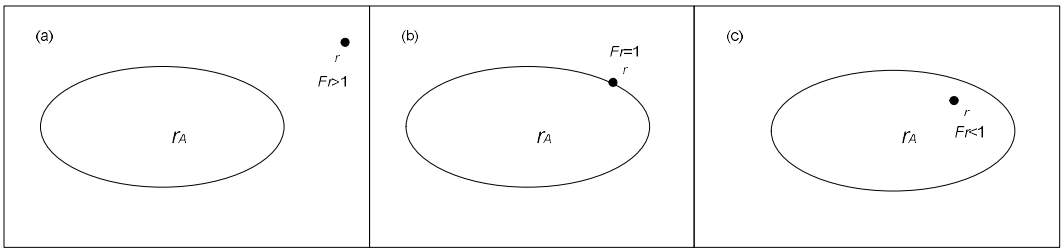



Figure 3.2.4 One point in space, r, and one ellipsoid with value of $\mathbf{F}_A$: (a) point outside ellipsoid, $\mathbf{F}_A > 1$; (b) point on ellipsoid, $\mathbf{F}_A = 1$; (c) point inside ellipsoid, $\mathbf{F}_A < 1$.

This means that if the contact potential value of point **r** is higher than one, **r** is outside the ellipsoid $A$; if contact potential value of point **r** is lower than one, **r** is inside the ellipsoid $A$; and finally if contact potential value equals one, **r** is on surface of the ellipsoid $A$.

Now consider the spatial relationship of two arbitrary ellipsoids (cells) $A$ and $B$. As discussed, their direction can be represented by matrices **U** and **V** respectively, in which the column vectors of each matrix represent the direction of semi-axis of each ellipsoid.

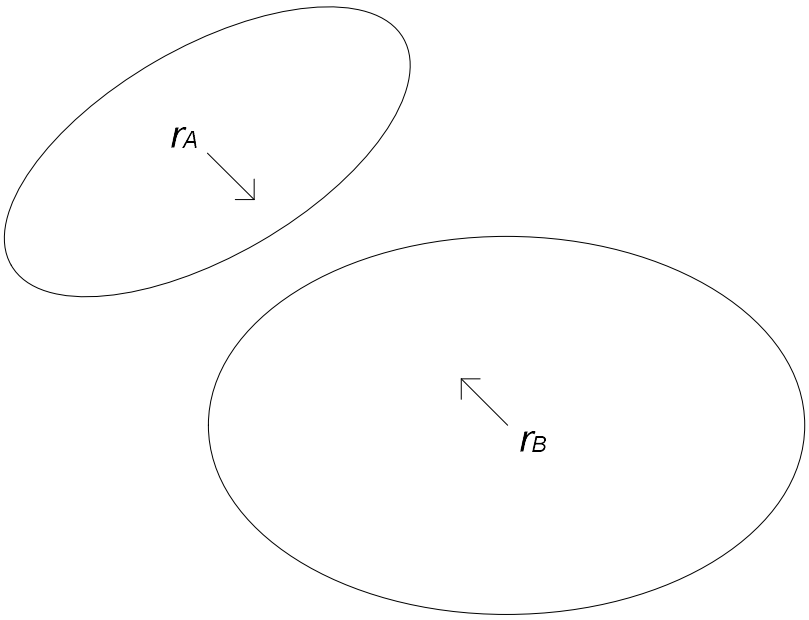



Figure 3.2.5 Two arbitrary ellipsoids and force on centre.

$$\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \tag{2}$$

$$\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \tag{3}$$

where $a_1$, $a_2$ and $a_3$ are the radii of ellipsoid $A$ and $b_1$, $b_2$ and $b_3$ are the radii of ellipsoid $B$ and they satisfy

$$0 < a_1 < a_2 < a_3 \tag{4}$$

$$0 < b_1 < b_2 < b_3 \tag{5}$$

Define each ellipsoid by the direction and length of all its semi-axis. Thus, $A$ and $B$ can be written as

$$\mathbf{A} = a_1^{-2}\mathbf{u}_1 \otimes \mathbf{u}_1 + a_2^{-2}\mathbf{u}_2 \otimes \mathbf{u}_2 + a_3^{-2}\mathbf{u}_3 \otimes \mathbf{u}_3 \tag{6}$$

$$\mathbf{B} = b_1^{-2}\mathbf{v}_1 \otimes \mathbf{v}_1 + b_2^{-2}\mathbf{v}_2 \otimes \mathbf{v}_2 + b_3^{-2}\mathbf{v}_3 \otimes \mathbf{v}_3 \tag{7}$$

The determinant calculation of both matrices $\mathbf{A}$ and $\mathbf{B}$ are larger than 0, thus it is easy to verify that the inverse matrices are

$$\mathbf{A}^{-1} = a_1^2\mathbf{u}_1 \otimes \mathbf{u}_1 + a_2^2\mathbf{u}_2 \otimes \mathbf{u}_2 + a_3^2\mathbf{u}_3 \otimes \mathbf{u}_3 \quad (8)$$

$$\mathbf{B}^{-1} = b_1^2\mathbf{v}_1 \otimes \mathbf{v}_1 + b_2^2\mathbf{v}_2 \otimes \mathbf{v}_2 + b_3^2\mathbf{v}_3 \otimes \mathbf{v}_3 \quad (9)$$

For an arbitrary point **r** in space, its potential to each ellipsoid can be written as

$$F_A(\mathbf{r}-\mathbf{r}_A) = (\mathbf{r}-\mathbf{r}_A)^T \mathbf{A}(\mathbf{r}-\mathbf{r}_A) \quad (10)$$

$$F_B(\mathbf{r}-\mathbf{r}_B) = (\mathbf{r}-\mathbf{r}_B)^T \mathbf{B}(\mathbf{r}-\mathbf{r}_B) \quad (11)$$

Next the potential of arbitrary point $\mathbf{r}$ is derived against ellipsoid $\mathbf{A}$ and $\mathbf{B}$ based on (13) and (14), thus the equation representing the potential between ellipsoid $\mathbf{A}$ and $\mathbf{B}$ should contain both $F_A(\mathbf{r}-\mathbf{r}_A)$ and $F_B(\mathbf{r}-\mathbf{r}_B)$. Similar to the equation 3.28 of [Allen, M. P. & Germano, 2002], a decimal $\lambda$ between 0 and 1 is introduced (the calculation of $\lambda$ is in Appendix A), then write the potential of point $\mathbf{r}$ against ellipsoid $\mathbf{A}$ and $\mathbf{B}$ as

$$F(\mathbf{r},\lambda) = \lambda F_A(\mathbf{r}-\mathbf{r}_A) + (1-\lambda)F_B(\mathbf{r}-\mathbf{r}_B) \quad (12)$$

In the equation $F(\mathbf{r},\lambda) = \lambda F_A(\mathbf{r}-\mathbf{r}_A) + (1-\lambda)F_B(\mathbf{r}-\mathbf{r}_B)$, the decimal $\lambda$ >=0, which means the both ellipsoids contribute to potential of point $\mathbf{r}$. If $\lambda$=0, it means that the point $\mathbf{r}$ is actually on center of one of ellipsoids. Note that there should be no fixed order of ellipsoid $\mathbf{A}$ and $\mathbf{B}$ in this equation， it can also be written as $F(\mathbf{r},\lambda) = \lambda F_B(\mathbf{r}-\mathbf{r}_B) + (1-\lambda)F_A(\mathbf{r}-\mathbf{r}_A)$ , because the potential of a fixed point should be fixed and is only relative to the geometry relationship between fixed point

and ellipsoid, not order of ellipsoid. Substitute (13), (14) into equation (15)

$$F(\mathbf{r},\lambda)=\lambda(\mathbf{r}-\mathbf{r}_A)^T\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+(1-\lambda)(\mathbf{r}-\mathbf{r}_B)^T\mathbf{B}(\mathbf{r}-\mathbf{r}_B) \tag{13}$$

For all $\lambda$ that satisfies $0\le\lambda\le1$, $F(\mathbf{r},\lambda)\ge0$, which means the value of potential of any point in space against two arbitrary ellipsoids should be equal or larger than zero. Also for any fixed $\lambda$, F has a unique minimum value has a function of r (J. W. Perram *et al*. 1985). Next this unique minimum value is estimate, then the condition that F reaches this value can be found.

To find out this minimum value, let $\dfrac{\partial F(\mathbf{r},\lambda)}{\partial\mathbf{r}}=0$, i.e.

$$2\lambda\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+2(1-\lambda)\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=0 \tag{14}$$

At the point where $\dfrac{\partial F(\mathbf{r},\lambda)}{\partial\mathbf{r}}=0$, function F reaches its minimum value. With the transformation focusing on ellipsoid A or B, two equations can be obtained, which are of $\mathbf{A}(\mathbf{r}-\mathbf{r}_A)$ and $\mathbf{B}(\mathbf{r}-\mathbf{r}_B)$ correspondingly

$$\mathbf{A}(\mathbf{r}-\mathbf{r}_A)=(1-\lambda)\left[\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}\right]^{-1}(\mathbf{r}_B-\mathbf{r}_A) \tag{15}$$

$$\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=\lambda\left[\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}\right]^{-1}(\mathbf{r}_A-\mathbf{r}_B) \tag{16}$$

It is obvious that they have common part $\left[\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}\right]^{-1}$. Similar to the form used in (J. W. Perram *et al*. 1996; H. Goldstein *et al*. 2001), this part is taken out and denoted as

$$\mathbf{G}(\lambda)=\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1} \tag{17}$$

Note that I consider $\lambda$ as a fixed value, then $\lambda$ is treated as a known value in function $\mathbf{G}(\lambda)=\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}$, and the geometry of ellipsoid A and B are also known. The the next step is to represent the vector $\mathbf{r}$ (the position of the arbitrary point) with $\lambda$, A and B. Equation $\mathbf{A}(\mathbf{r}-\mathbf{r}_A)=(1-\lambda)\left[\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}\right]^{-1}(\mathbf{r}_B-\mathbf{r}_A)$ and $\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=\lambda\left[\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}\right]^{-1}(\mathbf{r}_A-\mathbf{r}_B)$ can be transformed as

$$\mathbf{A}(\mathbf{r}-\mathbf{r}_A)=(1-\lambda)\mathbf{G}^{-1}(\lambda)(\mathbf{r}_B-\mathbf{r}_A) \tag{18}$$

$$\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=\lambda\mathbf{G}^{-1}(\lambda)(\mathbf{r}_A-\mathbf{r}_B) \tag{19}$$

And from $\mathbf{A}(\mathbf{r}-\mathbf{r}_A)=(1-\lambda)\mathbf{G}^{-1}(\lambda)(\mathbf{r}_B-\mathbf{r}_A)$

$$\mathbf{r}=\mathbf{r}_A+(1-\lambda)\mathbf{A}^{-1}\mathbf{G}^{-1}(\lambda)(\mathbf{r}_B-\mathbf{r}_A) \tag{20}$$

From $\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=\lambda\mathbf{G}^{-1}(\lambda)(\mathbf{r}_A-\mathbf{r}_B)$

$$\mathbf{r}=\mathbf{r}_B+\lambda\mathbf{B}^{-1}\mathbf{G}^{-1}(\lambda)(\mathbf{r}_A-\mathbf{r}_B) \tag{21}$$

Expand equation $2\lambda\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+2(1-\lambda)\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=0$ and get

$$\mathbf{r}=\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]^{-1}\left[\lambda\mathbf{A}\cdot\mathbf{r}_A+(1-\lambda)\mathbf{B}\cdot\mathbf{r}_B\right] \tag{22}$$

Take ellipsoid A and B as fixed variables, equation $\mathbf{r}=\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]^{-1}\left[\lambda\mathbf{A}\cdot\mathbf{r}_A+(1-\lambda)\mathbf{B}\cdot\mathbf{r}_B\right]$ tells that $\mathbf{r}$ is the function of $\lambda$. Remember that from equation $F(\mathbf{r},\lambda)=\lambda(\mathbf{r}-\mathbf{r}_A)^T\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+(1-\lambda)(\mathbf{r}-\mathbf{r}_B)^T\mathbf{B}(\mathbf{r}-\mathbf{r}_B)$, $F(\mathbf{r},\lambda)\geq 0$. Then the condition that $F(\mathbf{r},\lambda)$ reaches its minimum value is to be explored.

It is known that equation $2\lambda\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+2(1-\lambda)\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=0$ is derived with condition $\frac{\partial F(\mathbf{r},\lambda)}{\partial \mathbf{r}}=0$, where $F(\mathbf{r},\lambda)$ reaches its minimum value, and equation $\mathbf{r}=\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]^{-1}\left[\lambda\mathbf{A}\cdot\mathbf{r}_A+(1-\lambda)\mathbf{B}\cdot\mathbf{r}_B\right]$ is derived from equation $2\lambda\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+2(1-\lambda)\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=0$. Then equation $\mathbf{r}=\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]^{-1}\left[\lambda\mathbf{A}\cdot\mathbf{r}_A+(1-\lambda)\mathbf{B}\cdot\mathbf{r}_B\right]$ is the condition that makes $F(\mathbf{r},\lambda)$ reaches its minimum value. Now it is known that $F(\mathbf{r},\lambda)$ reaches its minimum value when $\lambda$ and $\mathbf{r}$ satisfy equation $\mathbf{r}=\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]^{-1}\left[\lambda\mathbf{A}\cdot\mathbf{r}_A+(1-\lambda)\mathbf{B}\cdot\mathbf{r}_B\right]$. In addition, this minimum value $F(\mathbf{r},\lambda)$ is the potential on point $\mathbf{r}$ against ellipsoid A and B.

Then it is needed to find the expression of $\lambda$ that makes $F(\mathbf{r},\lambda)=\lambda(\mathbf{r}-\mathbf{r}_A)^T\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+(1-\lambda)(\mathbf{r}-\mathbf{r}_B)^T\mathbf{B}(\mathbf{r}-\mathbf{r}_B)$ reaches its minimum value. Again note that $F(\mathbf{r},\lambda)$ is function of $\mathbf{r}$ and $\lambda$. Denote the minimum value of $F(\mathbf{r},\lambda)$ as follows

$$S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)=\min_{\mathbf{r}} F(\mathbf{r},\lambda) \tag{23}$$

As Figure 3.2.6 shows, with 2 intersectant ellipsoids centered at $\mathbf{r_A}$ and $\mathbf{r_B}$ and arbitrary point $\mathbf{r}$, there can be many possible value of $F(\mathbf{r},\lambda)$ at point $\mathbf{r}$ with different value of $\lambda$ (but $\lambda$ still should be between 0.0 and 1.0). $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)$ means the minimum value of all possible value of $F(\mathbf{r},\lambda)$ and is considered as potential of fixed point $\mathbf{r}$. From equation $F(\mathbf{r},\lambda)=\lambda(\mathbf{r}-\mathbf{r}_A)^T\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+(1-\lambda)(\mathbf{r}-\mathbf{r}_B)^T\mathbf{B}(\mathbf{r}-\mathbf{r}_B)$ it is known that $F(\mathbf{r},\lambda)$ is

a function of geometry information of ellipsoid A and B, $\mathbf{r}$ and $\lambda$. Although from equation $\mathbf{r}=\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]^{-1}\left[\lambda\mathbf{A}\cdot\mathbf{r}_A+(1-\lambda)\mathbf{B}\cdot\mathbf{r}_B\right]$ it is known that $\mathbf{r}$ changes with $\lambda$, the relationship between $\mathbf{r}$ and $\lambda$ is non-linear and it is too complex to discuss function $F(\mathbf{r},\lambda)$ with these two variables at the same time. In addition, for a fixed point $\mathbf{r}$ in space, the value of vector $\mathbf{r}$ is fixed. Thus I consider $\mathbf{r}$ as fixed value and do partial derivation of $\lambda$ against $F(\mathbf{r},\lambda)$, so that I can get to know about how $F(\mathbf{r},\lambda)$ changes with $\lambda$.

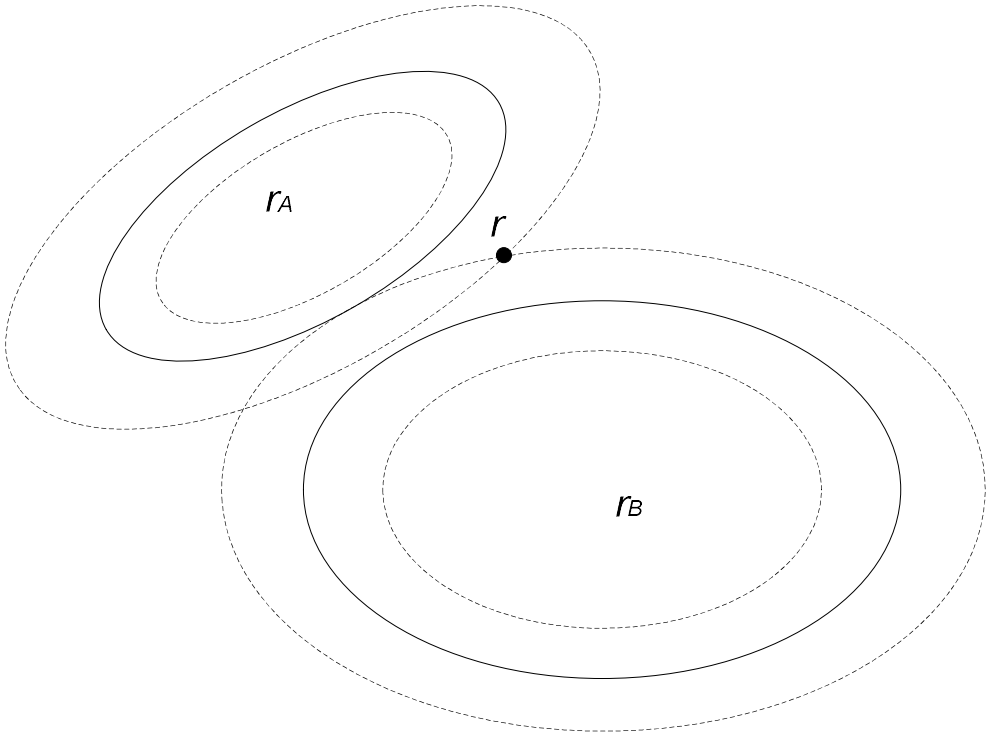



Figure 3.2.6 Two non-identical ellipsoids and one point in between that has the minimum value of $\mathbf{F}_A$

According to equation $\mathbf{A}(\mathbf{r}-\mathbf{r}_A)=(1-\lambda)\left[\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}\right]^{-1}(\mathbf{r}_B-\mathbf{r}_A)$, the potential on point $\mathbf{r}$ with relationship of ellipsoid A can be represented as

$$S_A(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)=F_A\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right]=\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right]^T\mathbf{A}\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right] \quad (24)$$

Similarly according to equation $\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=\lambda\left[\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}\right]^{-1}(\mathbf{r}_A-\mathbf{r}_B)$ , the potential on point $\mathbf{r}$ with relationship of ellipsoid B can be represented as

$$S_B(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)=F_B\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right]=\left[\mathbf{r}(\lambda)-\mathbf{r}_B\right]^T\mathbf{B}\left[\mathbf{r}(\lambda)-\mathbf{r}_B\right] \quad (25)$$

From equation $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)=\min_r F(\mathbf{r},\lambda)$ and

$$F(\mathbf{r},\lambda)=\lambda(\mathbf{r}-\mathbf{r}_A)^T\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+(1-\lambda)(\mathbf{r}-\mathbf{r}_B)^T\mathbf{B}(\mathbf{r}-\mathbf{r}_B)$$

$$\begin{aligned} S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B) &= F\left[\mathbf{r}(\lambda),\lambda\right] \\ &= \lambda F_A\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right]+(1-\lambda)F_B\left[\mathbf{r}(\lambda)-\mathbf{r}_B\right] \\ &= \lambda S_A(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)+(1-\lambda)S_B(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B) \end{aligned} \quad (26)$$

From equation $\mathbf{A}(\mathbf{r}-\mathbf{r}_A)=(1-\lambda)\left[\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}\right]^{-1}(\mathbf{r}_B-\mathbf{r}_A)$ and

$\mathbf{G}(\lambda)=\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}$

$$(\mathbf{r}-\mathbf{r}_A)^T=(1-\lambda)(\mathbf{r}_B-\mathbf{r}_A)^T\left[\mathbf{G}^{-1}(\lambda)\right]^T\left[\mathbf{A}^{-1}\right]^T \quad (27)$$

$$(\mathbf{r}-\mathbf{r}_B)^T=\lambda(\mathbf{r}_A-\mathbf{r}_B)^T\left[\mathbf{G}^{-1}(\lambda)\right]^T\left[\mathbf{B}^{-1}\right]^T \quad (28)$$

And $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)=\min_r F(\mathbf{r},\lambda)$ can be transformed to

$$S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)=\lambda(1-\lambda)(\mathbf{r}_A-\mathbf{r}_B)^T\mathbf{G}^{-1}(\lambda)(\mathbf{r}_A-\mathbf{r}_B) \quad (29)$$

Equation $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)=\lambda(1-\lambda)(\mathbf{r}_A-\mathbf{r}_B)^T\mathbf{G}^{-1}(\lambda)(\mathbf{r}_A-\mathbf{r}_B)$ is in similar form with (J. W. Perram *et al*. 1996). From conclusion in (Perram *et al*. 1996) it is known that $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)$ is non negative with $0\le\lambda\le1$ and zero at both end points. Now I test if $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)$ has single maximum value for any $\lambda$ in $0\le\lambda\le1$.

Note that with each fixed $\lambda$, there are many possible value of $F(\mathbf{r},\lambda)$ and $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)$ is the minimum value of $F(\mathbf{r},\lambda)$. The maximum value being discussed now is the highest value of $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)$ with all the possible value of $\lambda$ that satisfies $0 \le \lambda \le 1$, as Figure 3.2.7 shows.

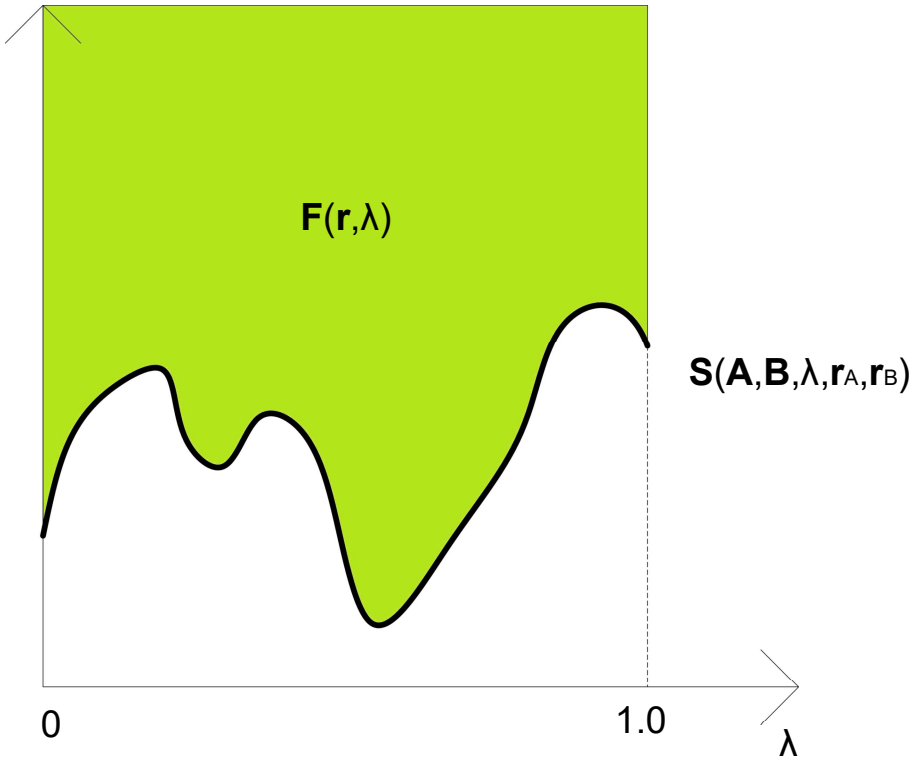



Figure 3.2.7 value of $F(\mathbf{r},\lambda)$ and $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)$; among all the possible value of $F(\mathbf{r},\lambda)$, $S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)=\min_r F(\mathbf{r},\lambda)$ is the minimum possible value of $F(\mathbf{r},\lambda)$ on $0 \le \lambda \le 1$

To study the feature of $F(\mathbf{r},\lambda)$ as a function, I start with feature of $\frac{d[\mathbf{r}(\lambda),\lambda]}{d\lambda}$. From equation $\mathbf{r}=\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]^{-1}\left[\lambda\mathbf{A}\cdot\mathbf{r}_A+(1-\lambda)\mathbf{B}\cdot\mathbf{r}_B\right]$ it is known that $\mathbf{r}$ is the function of $\lambda$, and $F(\mathbf{r},\lambda)=\lambda(\mathbf{r}-\mathbf{r}_A)^R\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+(1-\lambda)(\mathbf{r}-\mathbf{r}_B)^R\mathbf{B}(\mathbf{r}-\mathbf{r}_B)$. So equation $\frac{d[\mathbf{r}(\lambda),\lambda]}{d\lambda}$ turns to

$$\frac{dF\left[\mathbf{r}(\lambda),\lambda\right]}{d\lambda}=\frac{\partial F(\mathbf{r},\lambda)}{\partial\lambda}+\mathbf{r}'(\lambda)\cdot\frac{\partial F(\mathbf{r},\lambda)}{\partial\mathbf{r}}$$
$$=\left[(\mathbf{r}-\mathbf{r}_A)^T\mathbf{A}(\mathbf{r}-\mathbf{r}_A)-(\mathbf{r}-\mathbf{r}_B)^T\mathbf{B}(\mathbf{r}-\mathbf{r}_B)\right] \quad (30)$$
$$+\left[\lambda(\mathbf{r}-\mathbf{r}_A)^T\cdot\mathbf{A}+(1-\lambda)(\mathbf{r}-\mathbf{r}_B)^T\cdot\mathbf{B}\right]\cdot\mathbf{r}'(\lambda)$$

Using $2\lambda\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+2(1-\lambda)\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=0$

$$\frac{dF\left[\mathbf{r}(\lambda),\lambda\right]}{d\lambda}=\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right]^T\mathbf{A}\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right]-\left[\mathbf{r}(\lambda)-\mathbf{r}_B\right]^T\mathbf{B}\left[\mathbf{r}(\lambda)-\mathbf{r}_B\right] \quad (31)$$

From equation $\frac{dF\left[\mathbf{r}(\lambda),\lambda\right]}{d\lambda}=\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right]^T\mathbf{A}\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right]-\left[\mathbf{r}(\lambda)-\mathbf{r}_B\right]^T\mathbf{B}\left[\mathbf{r}(\lambda)-\mathbf{r}_B\right]$, it is known that $\frac{d[\mathbf{r}(\lambda),\lambda]}{d\lambda}$ can be larger or smaller than zero, which means $F(\mathbf{r},\lambda)$ increases with some value of $\lambda$ and decreases with some value of $\lambda$. This does not provide much useful information about existence of maximum value of $F(\mathbf{r},\lambda)$. Thus I derive the second derivative of $F(\mathbf{r},\lambda)$ to study $\frac{d[\mathbf{r}(\lambda),\lambda]}{d\lambda}$, i.e.

$$\frac{d^2F\left[\mathbf{r}(\lambda),\lambda\right]}{d\lambda^2}=2\left\{\left[\mathbf{r}(\lambda)-\mathbf{r}_A\right]^T\cdot\mathbf{A}-\left[\mathbf{r}(\lambda)-\mathbf{r}_B\right]^T\cdot\mathbf{B}\right\}\cdot\mathbf{r}'(\lambda) \quad (32)$$

After substituting equation (18), (30), and using (20), it can be transformed to

$$\frac{d^2F\left[\mathbf{r}(\lambda),\lambda\right]}{d\lambda^2}=2(\mathbf{r}_B-\mathbf{r}_A)^T\cdot\mathbf{G}^{-1}(\lambda)\cdot\mathbf{r}'(\lambda) \quad (33)$$

The term $\mathbf{r}'(\lambda)$ requires further expansion. At the position that $F(\mathbf{r},\lambda)$ reaches its minimum value, $\mathbf{r}(\lambda)$ satisfies $\frac{\partial F(\mathbf{r},\lambda)}{\partial\mathbf{r}}=0$. Then

$$\lambda\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+(1-\lambda)\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=0 \quad (34)$$

Integral equation $\lambda\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+(1-\lambda)\mathbf{B}(\mathbf{r}-\mathbf{r}_B)=0$ against $\lambda$

$$\frac{\partial\left(\lambda\mathbf{A}(\mathbf{r}-\mathbf{r}_A)+(1-\lambda)\mathbf{B}(\mathbf{r}-\mathbf{r}_B)\right)}{\partial\lambda}=0$$
$$\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]\mathbf{r}'(\lambda)=-\left[\mathbf{A}(\mathbf{r}-\mathbf{r}_A)-\mathbf{B}(\mathbf{r}-\mathbf{r}_B)\right] \quad (35)$$

Or

$$\mathbf{r}'(\lambda)=-\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]^{-1}\mathbf{G}^{-1}(\lambda)(\mathbf{r}_B-\mathbf{r}_A) \quad (36)$$

Substitute equation $\mathbf{r}'(\lambda)=-\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]^{-1}\mathbf{G}^{-1}(\lambda)(\mathbf{r}_B-\mathbf{r}_A)$ back to equation $\frac{d^2F\left[\mathbf{r}(\lambda),\lambda\right]}{d\lambda^2}=2(\mathbf{r}_B-\mathbf{r}_A)^T\cdot\mathbf{G}^{-1}(\lambda)\cdot\mathbf{r}'(\lambda)$, at position where $\lambda$ satisfies $\frac{\partial F(\mathbf{r},\lambda)}{\partial\mathbf{r}}=0$,

$$\frac{d^2F\left[\mathbf{r}(\lambda),\lambda\right]}{d\lambda^2}=-2(\mathbf{r}_B-\mathbf{r}_A)^T\left\{\mathbf{G}(\lambda)\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]\mathbf{G}(\lambda)\right\}^{-1}(\mathbf{r}_B-\mathbf{r}_A) \quad (37)$$

i.e.

$$\frac{d^2S(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B)}{d\lambda^2}=-2(\mathbf{r}_B-\mathbf{r}_A)^T\left\{\mathbf{G}(\lambda)\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]\mathbf{G}(\lambda)\right\}^{-1}(\mathbf{r}_B-\mathbf{r}_A) \quad (38)$$

Note that determinant calculation of both matrices $\mathbf{A}$ and $\mathbf{B}$ are larger than 0 (see equation $\mathbf{A}=a_1^{-2}\mathbf{u}_1\otimes\mathbf{u}_1+a_2^{-2}\mathbf{u}_2\otimes\mathbf{u}_2+a_3^{-2}\mathbf{u}_3\otimes\mathbf{u}_3$ and $\mathbf{B}=b_1^{-2}\mathbf{v}_1\otimes\mathbf{v}_1+b_2^{-2}\mathbf{v}_2\otimes\mathbf{v}_2+b_3^{-2}\mathbf{v}_3\otimes\mathbf{v}_3$ ) and $0\le\lambda\le1$, thus $1-\lambda\ge0$. Therefore $\lambda\mathbf{A}\ge0$ and $(1-\lambda)\mathbf{B}\ge0$. Because $\lambda\mathbf{A}$ and $(1-\lambda)\mathbf{B}$ cannot be zero at the same time, finally $\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]>0$, and $\mathbf{G}(\lambda)=\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}>0$, then $\left\{\mathbf{G}(\lambda)\left[\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\right]\mathbf{G}(\lambda)\right\}^{-1}>0$.

Thus $\left(\mathbf{r}_B - \mathbf{r}_A\right)^T \left\{\mathbf{G}\left(\lambda\right)\left[\lambda\mathbf{A}+\left(1-\lambda\right)\mathbf{B}\right]\mathbf{G}\left(\lambda\right)\right\}^{-1}\left(\mathbf{r}_B - \mathbf{r}_A\right) > 0$ , and finally $\frac{d^2 S\left(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B\right)}{d\lambda^2} < 0$, which means the value of $F\left(\mathbf{r},\lambda\right)$ increases more and more slowy as $\lambda$ increases, $F\left(\mathbf{r},\lambda\right)$ is a concave function. It implies the uniqueness of the maximum value of $F\left(\mathbf{r},\lambda\right)$ in the interval of $\lambda \in [0,1]$, $S\left(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B\right)$ is a concave function of $\lambda$, as Figure 3.2.8 shows.

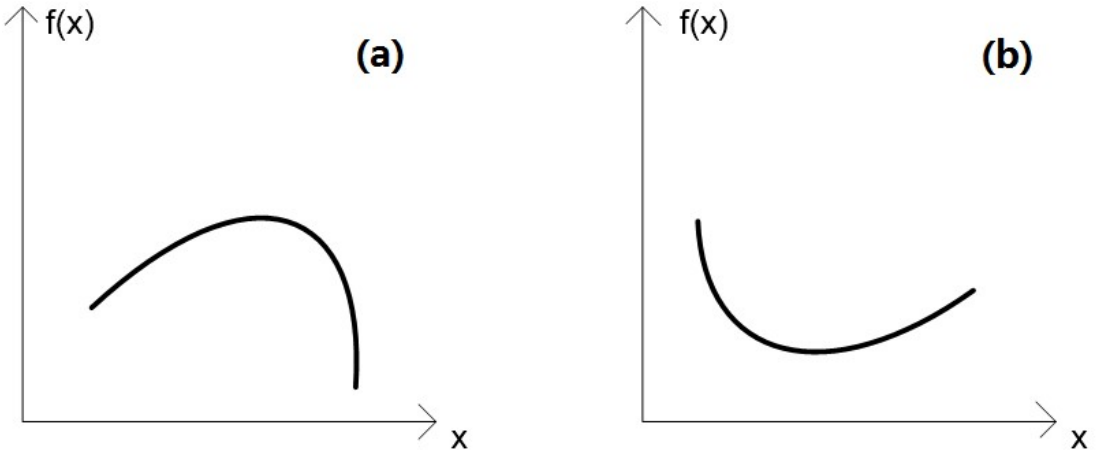



Figure 3.2.8 (a) concave function with a unique maximum value; (b) convex function with a unique minimum value

The value of $\lambda$ at this maximum value point is the value used to calculate pair potential of ellipsoid A and B. Denote this value as $\lambda_0$, then at this point $\lambda_0$ satisfies

$$\frac{dS\left(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B\right)}{d\lambda} = 0 \tag{39}$$

As discussed above, $\lambda_0$ is the value I look for.

According to equation $S\left(\mathbf{A},\mathbf{B},\lambda,\mathbf{r}_A,\mathbf{r}_B\right) = \lambda\left(1-\lambda\right)\left(\mathbf{r}_A - \mathbf{r}_B\right)^T \mathbf{G}^{-1}\left(\lambda\right)\left(\mathbf{r}_A - \mathbf{r}_B\right)$, denote

$$\mathbf{s} = \left(\mathbf{r}_A - \mathbf{r}_B\right) \tag{40}$$

Then

$$F(\lambda) = \lambda(1-\lambda)\mathbf{s}^T\mathbf{G}^{-1}(\lambda)\mathbf{s} \tag{41}$$

In summary, similar to (Perram *et al*. 1996), the New Perram & Wertheim ellipsoid contact potential may be expressed as

$$\Phi(\mathbf{A},\mathbf{B},\mathbf{s}) = F(\lambda_0) = \lambda_0(1-\lambda_0)\mathbf{s}^T\mathbf{G}^{-1}(\lambda_0)\mathbf{s} \tag{42}$$

In which $\Phi(\mathbf{A},\mathbf{B},\mathbf{s})$ is potential between ellipsoid A and B, and $\Phi(\mathbf{A},\mathbf{B},\mathbf{s})$ means it is a function of geometry information of A and B as well as the vector of centre of both ellipsoids.

For calculation of $\mathbf{G}(\lambda)$, first the value of $\lambda$ needs to be calculated. Because $\mathbf{G}(\lambda) = \lambda\mathbf{B}^{-1} + (1-\lambda)\mathbf{A}^{-1}$,

Define

$$\mathbf{M} = \mathbf{B}^{1/2}\mathbf{A}^{-1}\mathbf{B}^{1/2} \tag{43}$$

$$\mathbf{t} = \mathbf{B}^{1/2}\mathbf{s} \tag{44}$$

Assume

$$\mathbf{M} = \mathbf{W}\begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3^2 \end{pmatrix}\mathbf{W}^T = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{pmatrix}\begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3^2 \end{pmatrix}\begin{pmatrix} w_{11} & w_{21} & w_{31} \\ w_{12} & w_{22} & w_{32} \\ w_{13} & w_{23} & w_{33} \end{pmatrix} \tag{45}$$

Define

$$\mathbf{q} = \mathbf{W}^T\mathbf{t} \tag{46}$$

$$g_1(x)=\frac{q_1^2\left(m_1^2x^2-1\right)}{\left(m_1^2x+1\right)^2}+\frac{q_2^2\left(m_2^2x^2-1\right)}{\left(m_2^2x+1\right)^2}+\frac{q_3^2\left(m_3^2x^2-1\right)}{\left(m_3^2x+1\right)^2}=0 \tag{47}$$

$$\lambda=\frac{1}{1+x} \tag{48}$$

Note that

$$g_1'(x)=\frac{dg_1(x)}{dx}=2(1+x)\left[\frac{q_1^2m_1^2}{\left(m_1^2x+1\right)^3}+\frac{q_2^2m_2^2}{\left(m_2^2x+1\right)^3}+\frac{q_3^2m_3^2}{\left(m_3^2x+1\right)^3}\right] \tag{49}$$

$$g_1'(x)>0 \tag{50}$$

$$g_1(0)=-\left(q_1^2+q_2^2+q_3^2\right)<0 \tag{51}$$

$$g_1(+\infty)=\left(\frac{q_1}{m_1}\right)^2+\left(\frac{q_2}{m_2}\right)^2+\left(\frac{q_3}{m_3}\right)^2>0 \tag{52}$$

Newton-Raphson method (William H. Press *et al*. 2007) could be used to find root of Equation $g_1(x)=\frac{q_1^2\left(m_1^2x^2-1\right)}{\left(m_1^2x+1\right)^2}+\frac{q_2^2\left(m_2^2x^2-1\right)}{\left(m_2^2x+1\right)^2}+\frac{q_3^2\left(m_3^2x^2-1\right)}{\left(m_3^2x+1\right)^2}=0$ as a function, then from value of $x$ I can get value of $\lambda$ using equation (50). The next step is calculation of $\mathbf{G}^{-1}(\lambda)$. Let

$$\begin{aligned}\mathbf{H}(\lambda)&=\left[\lambda\mathbf{I}+(1-\lambda)\mathbf{B}^{1/2}\mathbf{A}^{-1}\mathbf{B}^{1/2}\right]\\&=\left[\lambda\mathbf{I}+(1-\lambda)\mathbf{M}\right]\end{aligned} \tag{53}$$

Then equation $\mathbf{G}(\lambda)=\lambda\mathbf{B}^{-1}+(1-\lambda)\mathbf{A}^{-1}$ turns to

$$\mathbf{G}(\lambda)=\mathbf{B}^{-1/2}\mathbf{H}(\lambda)\mathbf{B}^{-1/2} \tag{54}$$

Thus

$$\mathbf{G}^{-1}(\lambda)=\mathbf{B}^{1/2}\mathbf{H}^{-1}(\lambda)\mathbf{B}^{1/2} \tag{55}$$

Question reduce to how to calculate $\mathbf{H}^{-1}(\lambda)=\left[\lambda\mathbf{I}+(1-\lambda)\mathbf{M}\right]^{-1}$. Denote

$$\mathbf{M}=\mathbf{W}\begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3^2 \end{pmatrix}\mathbf{W}^T=\mathbf{W}\bar{\mathbf{M}}\mathbf{W}^T \tag{56}$$

Then

$$\begin{aligned}\mathbf{H}^{-1}(\lambda)&=\left[\lambda\mathbf{I}+(1-\lambda)\mathbf{W}\bar{\mathbf{M}}\mathbf{W}^T\right]^{-1}\\ &=\left[\mathbf{W}\left(\lambda\mathbf{I}+(1-\lambda)\bar{\mathbf{M}}\right)\mathbf{W}^T\right]^{-1}\\ &=\mathbf{W}\left[\lambda\mathbf{I}+(1-\lambda)\bar{\mathbf{M}}\right]^{-1}\mathbf{W}^T\end{aligned} \tag{57}$$

$$\mathbf{H}^{-1}(\lambda)=\mathbf{W}\begin{pmatrix} \lambda+(1-\lambda)m_1^2 & 0 & 0 \\ 0 & \lambda+(1-\lambda)m_2^2 & 0 \\ 0 & 0 & \lambda+(1-\lambda)m_3^2 \end{pmatrix}^{-1}\mathbf{W}^T \tag{58}$$

$$\mathbf{H}^{-1}(\lambda)=\mathbf{W}\begin{pmatrix} \dfrac{1}{\lambda+(1-\lambda)m_1^2} & 0 & 0 \\ 0 & \dfrac{1}{\lambda+(1-\lambda)m_2^2} & 0 \\ 0 & 0 & \dfrac{1}{\lambda+(1-\lambda)m_3^2} \end{pmatrix}\mathbf{W}^T \tag{59}$$

Finally

$$\mathbf{G}^{-1}(\lambda)=\mathbf{B}^{1/2}\mathbf{W}\begin{pmatrix} \dfrac{1}{\lambda+(1-\lambda)m_1^2} & 0 & 0 \\ 0 & \dfrac{1}{\lambda+(1-\lambda)m_2^2} & 0 \\ 0 & 0 & \dfrac{1}{\lambda+(1-\lambda)m_3^2} \end{pmatrix}\left[\mathbf{B}^{1/2}\mathbf{W}\right]^T \tag{60}$$

In this way the value of $\mathbf{G}^{-1}(\lambda)$ can be estimated with value of $\lambda$. This process is implemented by programming code and carried out in simulation.

According to equation $F_A(\mathbf{r}-\mathbf{r}_A)\begin{cases}<1 \ for \ \mathbf{r} \ \text{inside} \ A\\ =1 \ for \ \mathbf{r} \ \text{on the surface of} \ A:\\ >1 \ for \ \mathbf{r} \ \text{outside} \ A\end{cases}$

$$\Phi(\mathbf{A},\mathbf{B},\mathbf{s})\begin{cases}<1 \ for\ A\ and\ B\ overlapping\\ =1 \ for\ A\ and\ B\ externally\ tangent\\ >1 \ for\ A\ and\ B\ non\text{-}overlapping\end{cases} \quad (61)$$

According to the method to calculate $\mathbf{G}(\lambda)$, it is complex to draw graph for contact potential as a function. As a simplified condition, for two identical ellipsoidal agents with three semi-axes equals 5, 2 and 2 and in the same direction, the value of potential versus distance is shown in Figure 3.2.9.

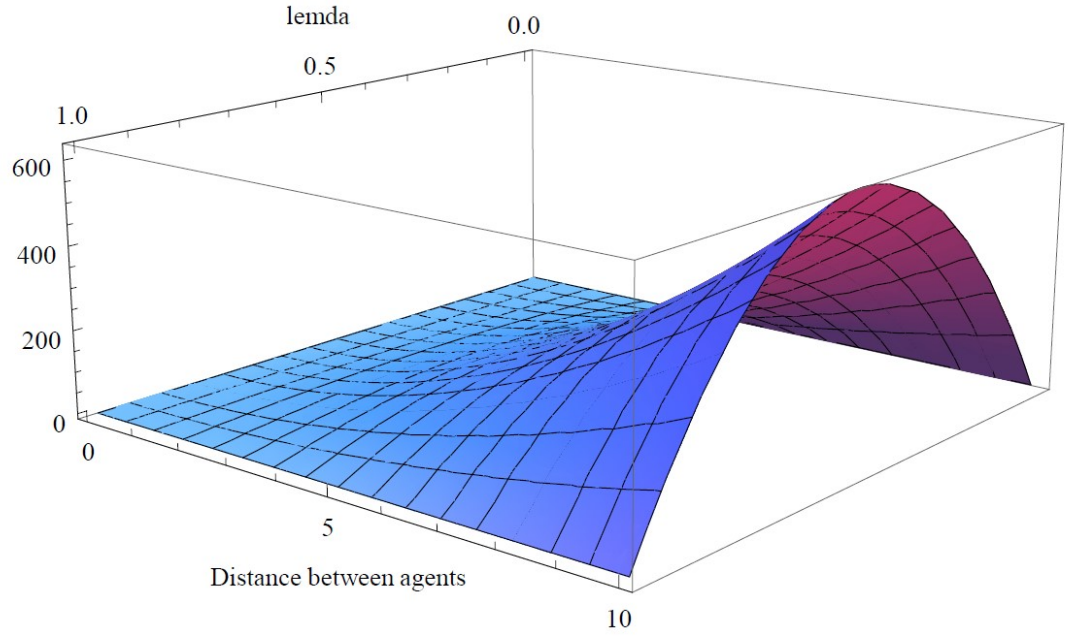



Figure 3.2.9 potential vs. distance

## 3.3 Adhesion force

"The dynamic balance between cell adhesion and cell movement is arguably the most important universal feature of adhesion in development" (Bowers-Morrow *et al.* 2004). In addition, the type and strength of different adhesion forces plays a significant role in cell sorting (Eirikur Palsson, 2008). The adhesion force likewise

plays an important role in agent motility in my model. The velocity of agent movement is controlled by the combination of adhesion force and resistance force, while the direction of agent movement is controlled by adhesion force alone. When one agent comes into contact with other agents, the adhesion force balances with the contact force. The point at which the adhesion force and contact force balance represents the elasticity of the agent. The necessity of adhesion force in such modelled systems is discussed in section 2.1.6.

The adhesion force depends on potential. First, I discuss how to generate force from potential in general. In the model, the potential is a scalar quantity that has value but no direction: potential represents some form of energy. According to classical mechanics, force does work to move an object from one point to another point in space. In general, a force is represented by a vector in space, in the form of $\mathbf{F}(x \quad y \quad z)$. The size of force is $|\mathbf{F}|$ and the direction of force is $\frac{\mathbf{F}}{|\mathbf{F}|}$. Let the start point be $\mathbf{P}_a$ and end point be $\mathbf{P}_b$, and both $\mathbf{P}_a$ and $\mathbf{P}_b$ are vectors in space. The energy $E$ equals the work done by force: $E = \int_{\mathbf{P}_a}^{\mathbf{P}_b} \mathbf{F}$. Like potential, the energy has no direction, and is a scalar quantity. Note that the value of $E$ can be positive or negative, meaning the force $\mathbf{F}$ does positive work or negative work respectively. If the energy $E$ and start and end point $\mathbf{P}_a$ and $\mathbf{P}_b$ are known, to calculate $\mathbf{F}$ I need to calculate the partial differential of $E$ by $\mathbf{P}_a - \mathbf{P}_b$, i.e. $\mathbf{F} = \frac{\partial E}{\partial(\mathbf{P}_a - \mathbf{P}_b)}$.

Here, as I assume the mass of agents are evenly distributed, the force can be considered to work on the centre point of the ellipsoid. The position of the ellipsoid is represented by the position of its centre. Specifically, in my model, for two ellipsoids, the energy $E$ equals the work done by force $\mathbf{F}$ to move an ellipsoid from its position to the position of the other ellipsoid. Similar to the previous deduction, in my model $\mathbf{F}=\frac{\partial E}{\partial(\mathbf{A}-\mathbf{B})}$, where $\mathbf{A}$ and $\mathbf{B}$ are position of centre of the two ellipsoids.

According to 2.4.4, the Metropolis algorithms (Metropolis *et al*. 1959) is are a group of methods in which the spatial relationship between cells is considered to have a particular energy, and cells are considered to move to minimize the energy of the system (Nicholas Metropolis *et al*. 1959). Here the Metropolis algorithm is used to generate energy from potential. After that the energy is partially derivated with respect to displacement to determine force. Torque is also generated from energy and the translation from energy to torque is discussed in Section 3.6. In order to simplify the problem, I begin by treating adhesion force for two identical spheres, and then generalise the conclusion to two arbitrary spheres, and further that generalisation to two arbitrary ellipsoids.

In the case of two identical spherical particles $\mathbf{A}$ and $\mathbf{B}$ with radius of R, because the sphere has no direction, according to equation $\mathbf{U}=(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3)=\begin{pmatrix}u_{11}&u_{12}&u_{13}\\u_{21}&u_{22}&u_{23}\\u_{31}&u_{32}&u_{33}\end{pmatrix}$, its direction matrix is the unit matrix, and

$$\mathbf{A}=\mathbf{U}\bar{\mathbf{A}}\mathbf{U}^{T}=\begin{pmatrix}u_{11}&u_{12}&u_{13}\\u_{21}&u_{22}&u_{23}\\u_{31}&u_{32}&u_{33}\end{pmatrix}\begin{pmatrix}\frac{1}{a_1^2}&0&0\\0&\frac{1}{a_2^2}&0\\0&0&\frac{1}{a_3^2}\end{pmatrix}\begin{pmatrix}u_{11}&u_{21}&u_{31}\\u_{12}&u_{22}&u_{32}\\u_{13}&u_{23}&u_{33}\end{pmatrix}$$ becomes

$$A=B=\begin{pmatrix}\frac{1}{R^2}&0&0\\0&\frac{1}{R^2}&0\\0&0&\frac{1}{R^2}\end{pmatrix} \tag{1}$$

According to equation $\Phi(\mathbf{A},\mathbf{B},\mathbf{s})=F(\lambda_0)=\lambda_0(1-\lambda_0)\mathbf{s}^T\mathbf{G}^{-1}(\lambda_0)\mathbf{s}$ the potential is therefore

$$\begin{aligned}\Phi(\mathbf{A},\mathbf{B},\mathbf{s})&=f(\lambda_0)=\lambda_0(1-\lambda_0)\mathbf{s}^T\mathbf{G}^{-1}(\lambda_0)\mathbf{s}\\&=\frac{(x_A-x_B)^2+(y_A-y_B)^2+(z_A-z_B)^2}{(2R)^2}\end{aligned} \tag{2}$$

In which $\mathbf{r_A}=(x_\mathbf{A},y_\mathbf{A},z_\mathbf{A})$ and $\mathbf{r_B}=(x_\mathbf{B},y_\mathbf{B},z_\mathbf{B})$ are centre of the two spherical particles $\mathbf{A}$ and $\mathbf{B}$.

In (Dallon and Othmer, 2004) and (Galle *et al*. 2005), the adhesion energy is modelled in the form of

$$W=\varepsilon A \tag{3}$$

In which $\varepsilon$ is a proportional constant, $A$ is the contact area of the two spheres, which is proportional to the distance between two spheres. In my model potential is proportional to square of distance between two ellipsoids, thus to generalise the equation for ellipsoids, I use square root of potential to replace the contact area term $A$

in equation $W = \varepsilon A$ and get

$$W^{ad} = \varepsilon\sqrt{\Phi(\mathbf{A},\mathbf{B},\mathbf{s})} \tag{4}$$

Note that the potential $\Phi$ is written in the form of $\Phi(\mathbf{A},\mathbf{B},\mathbf{s})$ to represent that the potential is a function of the geometry of A and B, and the distance vector between them. Since a sphere can be considered as a special type of ellipsoid with three semi-axes of same length, the potential formula can be shaped to work for spheres.

Substitute equation $\Phi(\mathbf{A},\mathbf{B},\mathbf{s}) = f(\lambda_0) = \lambda_0(1-\lambda_0)\mathbf{s}^{\mathrm{T}}\mathbf{G}^{-1}(\lambda_0)\mathbf{s} = \dfrac{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}{(2R)^2}$ to equation $W^{ad} = \varepsilon\sqrt{\Phi(\mathbf{A},\mathbf{B},\mathbf{s})}$ then

$$W^{Ad} = \varepsilon\sqrt{\frac{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}{(2R)^2}} \tag{5}$$

Then as discussed in Section 2.4.4, the value of force equals the partial differential of energy against the displacement. Thus the adhesion force is

$$F_A^{Ad} = -\frac{\partial W^{Ad}}{\partial r_A} = -\frac{\varepsilon}{2R}\frac{\partial\sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}}{\partial r_A}$$
$$= -\frac{\varepsilon}{4R}\frac{(x_A - x_B,\quad y_A - y_B,\quad z_A - z_B)}{\sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}} \tag{6}$$

This gives

$$\left|\mathbf{F}_A^{Ad}\right| = \frac{\varepsilon}{4R} = const \tag{7}$$

This means that the size of the adhesion force for two identical spheres is a constant

value. This may be counterintuitive, as one might think that the closer two particles are, the larger the adhesion force should be. However this assumption may be driven by the phenomenon that the movement speed of an object increases over time, which only indicates that direction of the force does not change. A constant force can, of course, explain the acceleration of an object.

This is the form of adhesion energy between two identical spheres as special case of ellipsoids in my model. The next step is to generalise the conclusion for two arbitrary spheres. In the case of two non-identical spherical particles, let $R_A$ be radius of sphere A and $R_B$ be radius of sphere B, similar to equation

$$A = B = \begin{pmatrix} \frac{1}{R^2} & 0 & 0 \\ 0 & \frac{1}{R^2} & 0 \\ 0 & 0 & \frac{1}{R^2} \end{pmatrix},$$

$$\mathbf{A} = \begin{pmatrix} \frac{1}{R_A^2} & 0 & 0 \\ 0 & R_A^2 & 0 \\ 0 & 0 & R_A^2 \end{pmatrix} \tag{8}$$

$$\mathbf{B} = \begin{pmatrix} \frac{1}{R_B^2} & 0 & 0 \\ 0 & R_B^2 & 0 \\ 0 & 0 & R_B^2 \end{pmatrix} \tag{9}$$

Substitute $\mathbf{A}=\begin{pmatrix} \frac{1}{R_A^2} & 0 & 0 \\ 0 & R_A^2 & 0 \\ 0 & 0 & R_A^2 \end{pmatrix}$ and $\mathbf{B}=\begin{pmatrix} \frac{1}{R_B^2} & 0 & 0 \\ 0 & R_B^2 & 0 \\ 0 & 0 & R_B^2 \end{pmatrix}$ to equation (46),

$$\Phi(\mathbf{A},\mathbf{B},\mathbf{s})=\frac{1}{(R_A+R_B)^2}\left[(x_A-x_B)^2+(y_A-y_B)^2+(z_A-z_B)^2\right] \quad (10)$$

Similar to equation $W^{Ad}=\varepsilon\sqrt{\frac{(x_A-x_B)^2+(y_A-y_B)^2+(z_A-z_B)^2}{(2R)^2}}$,

$$W^{Ad}=\varepsilon\sqrt{\frac{(x_A-x_B)^2+(y_A-y_B)^2+(z_A-z_B)^2}{(R_A+R_B)^2}} \quad (11)$$

Similar to $F_A^{Ad}=-\frac{\partial W^{Ad}}{\partial r_A}=-\frac{\varepsilon}{2R}\frac{\partial\sqrt{(x_A-x_B)^2+(y_A-y_B)^2+(z_A-z_B)^2}}{\partial r_A}=-\frac{\varepsilon}{4R}\frac{(x_A-x_B,y_A-y_B,z_A-z_B)}{\sqrt{(x_A-x_B)^2+(y_A-y_B)^2+(z_A-z_B)^2}}$, adhesion force

$$\mathbf{F}_A^{Ad}=-\frac{\partial W^{Ad}}{\partial \mathbf{r}_A}=-\frac{\varepsilon}{R_A+R_B}\frac{\partial\sqrt{(x_A-x_B)^2+(y_A-y_B)^2+(z_A-z_B)^2}}{\partial \mathbf{r}_A}=-\frac{\varepsilon}{2(R_A+R_B)}\frac{\{x_A-x_B,y_A-y_B,z_A-z_B\}}{\sqrt{(x_A-x_B)^2+(y_A-y_B)^2+(z_A-z_B)^2}} \quad (12)$$

This gives

$$\left|\mathbf{F}_A^{Ad}\right|=\frac{\varepsilon}{2(R_A+R_B)}=const \quad (13)$$

Which is in similar form of equation $\left|\mathbf{F}_A^{Ad}\right|=\frac{\varepsilon}{4R}=const$. Again the adhesion force is

constant. As a generalisation I assume for pair of arbitrary ellipsoidal particles:

$$W^{Ad} = \varepsilon\sqrt{\Phi(\mathbf{A},\mathbf{B},\mathbf{s})} \tag{14}$$

Then the adhesion force should be in form of

$$\mathbf{F}^{Ad} = \frac{\partial W^{Ad}}{\partial \mathbf{s}} = \frac{\partial W^{Ad}}{\partial \Phi}\cdot\frac{\partial \Phi}{\partial \mathbf{s}} = -2\lambda_0(1-\lambda_0)\frac{\partial W^{Ad}}{\partial \Phi}\mathbf{G}^{-1}(\lambda_0)\mathbf{s} \tag{15}$$

From $W^{Ad} = \varepsilon\sqrt{\Phi(\mathbf{A},\mathbf{B},\mathbf{s})}$

$$\frac{\partial W^{Ad}}{\partial \Phi} = \frac{\partial\left(\varepsilon\sqrt{\Phi}\right)}{\partial \Phi} = \frac{1}{2}\varepsilon\Phi^{-\frac{1}{2}} \tag{16}$$

Thus the adhesion force

$$\mathbf{F}^{Ad} = -\varepsilon\lambda_0(1-\lambda_0)\Phi^{-\frac{1}{2}}\mathbf{G}^{-1}(\lambda_0)\mathbf{s} \tag{17}$$

From $\mathbf{F}^{Ad} = -\varepsilon\lambda_0(1-\lambda_0)\Phi^{-\frac{1}{2}}\mathbf{G}^{-1}(\lambda_0)\mathbf{s}$, the adhesion force of two arbitrary ellipsoids is not constant, and its direction is always from centre of the reference ellipsoid to the centre of the other ellipsoid.

## 3.4 Contact force

As the adhesion force tends to make cells stick together: if there is no other force to balance it, all cells will end up completely overlapping. Contact force is the main interaction factor that prevents ellipsoid cells from overlapping. Similar to adhesion force, contact force is also calculated from energy, which is calculated from potential by using the Metropolis algorithm.

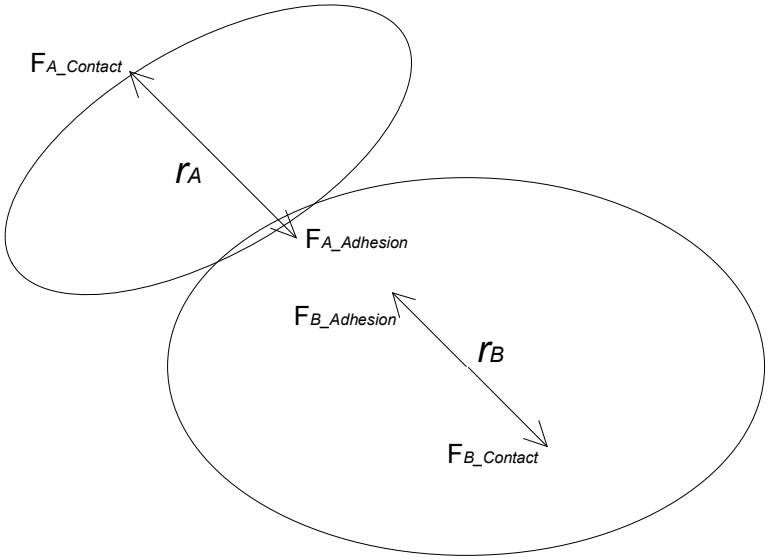



Figure 3.4.1 Contact force and adhesion force on two non-identical ellipsoids

The Lennard-Jones Potential is the most commonly used form for two-object interaction, which can be written as the following form:

$$V(r) = 4\varepsilon\left[\left(\frac{\sigma^2}{r^2}\right)^6 - \left(\frac{\sigma^2}{r^2}\right)^3\right] \tag{1}$$

The term $\frac{r^2}{\sigma^2}$ represents the distance between a pair of spheres. Since in my model the shape of the cell is ellipsoid, I cannot use $\frac{r^2}{\sigma^2}$ directly. To represent the spatial relationship of a pair of ellipsoids, similar to (Perram *et al*. 1996), I may use potential to replace $\frac{r^2}{\sigma^2}$. I then get contact energy

$$U = 4\varepsilon\left[\Phi^{-6}(A,B,s) - \Phi^{-3}(A,B,s)\right] \tag{2}$$

$$U = 4\varepsilon\left[\frac{1}{\Phi^6(\mathbf{A},\mathbf{B},\mathbf{r}_A,\mathbf{r}_B)} - \frac{1}{\Phi^3(\mathbf{A},\mathbf{B},\mathbf{r}_A,\mathbf{r}_B)}\right] \tag{3}$$

Where $U$ is the contact energy, $\Phi(\mathbf{A},\mathbf{B},\mathbf{r}_A,\mathbf{r}_B)$ is the potential, $\mathbf{r}_A$ and $\mathbf{r}_B$ are vectors that represent position of ellipsoid $A$ and $B$ respectively. According to Section 3.3, contact force is $\mathbf{F}=\frac{\partial U}{\partial(\mathbf{r}_A-\mathbf{r}_B)}$. Thus the ellipsoid contact force can be written in the following form.

$$\begin{aligned} F(r) &= \frac{\partial U}{\partial(\mathbf{r}_A-\mathbf{r}_B)} \\ &= \frac{\partial\left\{4\varepsilon\left[\frac{1}{\Phi^6(\mathbf{A},\mathbf{B},\mathbf{r}_A,\mathbf{r}_B)}-\frac{1}{\Phi^3(\mathbf{A},\mathbf{B},\mathbf{r}_A,\mathbf{r}_B)}\right]\right\}}{\partial(\mathbf{r}_A-\mathbf{r}_B)} \end{aligned} \tag{4}$$

Note that $\mathbf{s}=(\mathbf{r}_A-\mathbf{r}_B)$, then $F(r)=\frac{\partial U}{\partial(\mathbf{r}_A-\mathbf{r}_B)}$ $=\frac{\partial\left\{4\varepsilon\left[\frac{1}{\Phi^6(\mathbf{A},\mathbf{B},\mathbf{r}_A,\mathbf{r}_B)}-\frac{1}{\Phi^3(\mathbf{A},\mathbf{B},\mathbf{r}_A,\mathbf{r}_B)}\right]\right\}}{\partial(\mathbf{r}_A-\mathbf{r}_B)}$ turns to

$$\begin{aligned} F(r) &= -4\varepsilon\left[-6\Phi^{-7}(A,B,s)+3\Phi^{-4}(A,B,s)\right]\frac{\partial\Phi(A,B,s)}{\partial s} \\ &= -4\varepsilon\left[-6\Phi^{-7}(A,B,s)+3\Phi^{-4}(A,B,s)\right]\cdot\left[2\lambda_0(1-\lambda_0)G^{-1}(\lambda_0)s\right] \end{aligned} \tag{5}$$

Then I run an early stage simulation including contact and adhesion force and regularly save output data includes cell position, direction, velocity and angular velocity. A third party visualisation program (Simon Coakley, 2007. “X-agents 3D Visualisation”) reads these data and draws them on screen using an OpenGL API.

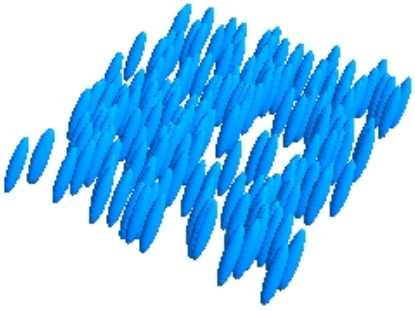


Figure 3.4.2 Starting condition of early simulation: 200 agents are randomly placed on the same substrate plane.

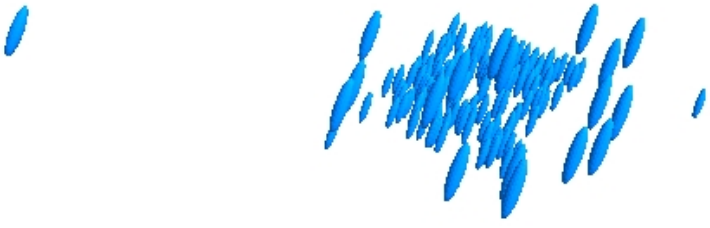


Figure 3.4.3 After a few loops of running of simulation, some agents have already

moved relatively far away from others.

Figure 3.4.4 After 10 seconds (simulation time), most of the agents have moved beyond the observation area.

Comparing Figure 3.4.2, Figure 3.4.3 and Figure 3.4.4, I can see the movement of agents is not realistic, in which the agents are forced apart from each other by a relatively large force. This phenomenon is caused by equation $F(r) = -4\varepsilon\left[-6\Phi^{-7}(A,B,s) + 3\Phi^{-4}(A,B,s)\right]\frac{\partial\Phi(A,B,s)}{\partial s}$ $= -4\varepsilon\left[-6\Phi^{-7}(A,B,s) + 3\Phi^{-4}(A,B,s)\right]\cdot\left[2\lambda_0(1-\lambda_0)G^{-1}(\lambda_0)s\right]$, which contains $\Phi^{-7}(A,B,s)$, a term that contains high power of displacement **s**. $\Phi^{-7}(A,B,s)$ makes the model dynamics highly sensitive to the value of potential, especially in the case when the value of $\Phi(A,B,s)$ is relatively small, which happens when two cells are close to each other. As shown in Figure 3.4.5, the energy dramatically increases

while potential decreases. When two agents are very close to each other, a small decrease of potential results a rapid increase of contact force which is much larger than the adhesion force, and therefore forces the agent apart at high speed. Clearly this is improper behaviour. The Lennard-Jones potential is used in (Michael P. Allen, 2004) only to deal with non-contact particle interaction; another formula that contains square of displacement is introduced for particle contact. I also need to switch to a formula with a lower power to avoid the problem. In (Ramis-Conde *et al*. 2009), (Ramis-Conde *et al*. 2008) and (Galle *et al*. 2008), the Hertz model is used to describe the physical interaction among cells.

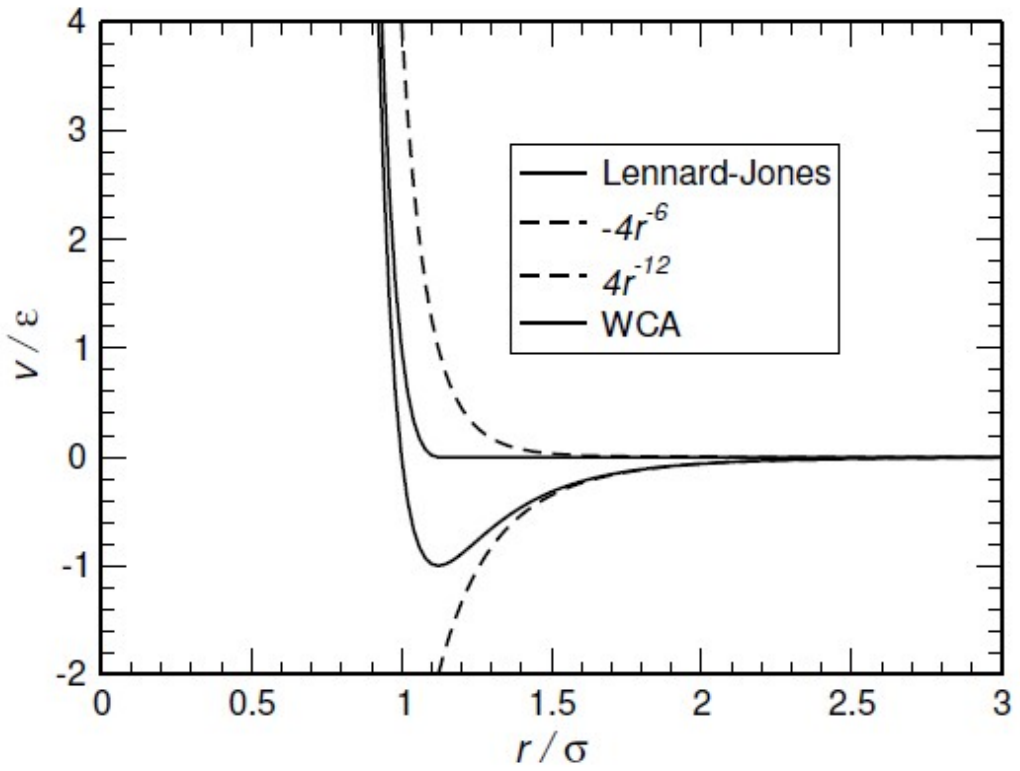



Figure 3.4.5 Energy dramatically increases while potential decreases (Michael P. Allen, 2004)

The Hertz model was introduced by Heinrich Hertz in 1882 to solve the contact force between two elastic objects (H. R. Hertz, 1882). The Hertz model is used to model the attractive and repulsive energy among elastic spheroidal cells. In the Hertz model, energy W between two spheroidal agents **A** and **B** can be represented as follows

$$W = \frac{2Ka^5}{5R_m^2} \tag{6}$$

In which

$$a = \sqrt{R_m \delta} \tag{7}$$

$$\delta = R_A + R_B - \left\| \mathbf{x}_A - \mathbf{x}_B \right\| \tag{8}$$

$$K = \frac{4}{3\pi\left(k_1 + k_2\right)} \tag{9}$$

$$\begin{cases} k_1 = \dfrac{1-\nu_1^2}{\pi E_1} \\ k_2 = \dfrac{1-\nu_2^2}{\pi E_2} \end{cases} \tag{10}$$

$$R_m = \frac{R_1 R_2}{R_1 + R_2} \tag{11}$$

In which $R_1$ and $R_2$ are radii of the two spheres; $\nu_1$ and $\nu_2$ are Poisson's ratios of **A** and **B** respectively, and $E_1$ and $E_2$ are Young's moduli of A and B respectively; $\delta$ is level of deformation of the two spheres. From $W = \frac{2Ka^5}{5R_m^2}$, $K = \frac{4}{3\pi\left(k_1 + k_2\right)}$, and $\begin{cases} k_1 = \dfrac{1-\nu_1^2}{\pi E_1} \\ k_2 = \dfrac{1-\nu_2^2}{\pi E_2} \end{cases}$, it is known that the energy W is proportional to Young's moduli of both spheres A and B, and is proportional to Poisson's ratios of A and B.

Also from $W = \frac{2Ka^5}{5R_m^2}$, $a = \sqrt{R_m \delta}$, and $\delta = R_A + R_B - \left\| \mathbf{x}_A - \mathbf{x}_B \right\|$ it is known that energy W is proportional to level of deformation. A and from equation $a = \sqrt{R_m \delta}$ and $\delta = R_A + R_B - \left\| \mathbf{x}_A - \mathbf{x}_B \right\|$, $a$ is proportional to square root of deformation, and then from equation $W = \frac{2Ka^5}{5R_m^2}$, the energy is proportional to 2.5 power of level of deformation, comparing to the seventh power in Lennard-Jones Potential, Hertz model should have smooth curve. Thus I decide to use Hertz model instead of Lennard-Jones Potential.

However the Hertz formula is not suitable for direct use with ellipsoidal agents as in my model, because an ellipsoid does not have radius R but three semi-axes which may not be equal to each other. Thus a generalisation is required. The goal of my generalisation is to replace all the terms that contain $R_1$ and $R_2$ in the formula, so that when substituting the same values for the lengths of the three semi-axes of the ellipsoid into a generalised Hertz formula, the generalised Hertz formula can revert to the original Hertz formula.

As per the approach of generalising the formula for contact potential, I begin with consideration of the Hertz formula for an ellipsoid and an arbitrary point in space, and then I apply the formula for two ellipsoids.

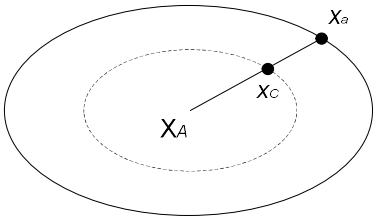



Figure 3.4.5 Ellipsoid $\mathbf{A}$ and two points $\mathbf{x}_a$ and $\mathbf{x}_c$. Point $\mathbf{x}_a$ on the surface of an ellipsoid and a point $\mathbf{x}_c$ inside an ellipsoid, and the two points are on the same line. The dotted line represents the potential surface on which point $\mathbf{x}_c$ lies.

For an ellipsoid $\mathbf{X}_A$ (here I use the position vector of the centre point of the ellipsoid to represent it) in space and one point $\mathbf{x}_c$ inside the ellipsoid. The contact force applies when two ellipsoids make contact with each other, thus I set point $\mathbf{x}_c$ inside the ellipsoid. By connecting point $\mathbf{x}_c$ and $\mathbf{X}_A$ and extending the line, I can find the point of intersection $\mathbf{x}_a$ on surface of ellipsoid. Similar to equation $F_A(\mathbf{r}-\mathbf{r}_A)=(\mathbf{r}-\mathbf{r}_A)^T\mathbf{A}(\mathbf{r}-\mathbf{r}_A)$ and $F_B(\mathbf{r}-\mathbf{r}_B)=(\mathbf{r}-\mathbf{r}_B)^T\mathbf{B}(\mathbf{r}-\mathbf{r}_B)$, the potential of point $\mathbf{x}_a$ to ellipsoid $\mathbf{X}_A$ can be written as follows

$$\mathbf{\Phi}_a=(\mathbf{x}_a-\mathbf{X}_A)^T\mathbf{A}(\mathbf{x}_a-\mathbf{X}_A) \tag{12}$$

where $\mathbf{\Phi}_a$ is the potential and A is the geometry matrix of ellipsoid $\mathbf{X}_A$.

Similarly for $\mathbf{x}_c$,

$$\mathbf{\Phi}_c=(\mathbf{x}_c-\mathbf{X}_A)^T\mathbf{A}(\mathbf{x}_c-\mathbf{X}_A) \tag{13}$$

Geometrically vector $(\mathbf{x}_c-\mathbf{X}_A)$ and $(\mathbf{x}_a-\mathbf{X}_A)$ are of same direction, and so I assume they satisfy

$$(\mathbf{x}_c-\mathbf{X}_A)=\mu(\mathbf{x}_a-\mathbf{X}_A) \tag{14}$$

Where $\mu$ is a real number and $0 < \mu < 1$. Thus

$$\mathbf{\Phi}_c = \mu^2 \left(\mathbf{x}_a - \mathbf{X}_A\right)^T \mathbf{A}\left(\mathbf{x}_a - \mathbf{X}_A\right) = \mu^2 \mathbf{\Phi}_a \tag{15}$$

Note that $\mathbf{x}_a$ is on the surface of the ellipsoid A, thus $\mathbf{\Phi}_a = 1$, i.e. $\mathbf{\Phi}_c = \mu^2$. So that

$$\mu = \sqrt{\mathbf{\Phi}_c} \tag{16}$$

Substitute $\mu = \sqrt{\mathbf{\Phi}_c}$ to equation $\left(\mathbf{x}_c - \mathbf{X}_A\right) = \mu\left(\mathbf{x}_a - \mathbf{X}_A\right)$,

$$\frac{\left(\mathbf{x}_c - \mathbf{x}_A\right)}{\sqrt{\mathbf{\Phi}_c}} = \left(\mathbf{x}_a - \mathbf{x}_A\right) \tag{17}$$

Thus

$$\mathbf{x}_a = \mathbf{x}_A + \frac{\mathbf{x}_c - \mathbf{x}_A}{\sqrt{\mathbf{\Phi}_c}} \tag{18}$$

Now I consider two arbitrary ellipsoids $\mathbf{A}$ and $\mathbf{B}$ in contact with each other, as Figure 3.4.6 shows. $\mathbf{X}_A$ and $\mathbf{X}_B$ are position vectors of centres of ellipsoid $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{x}_c$ is a point inside the ellipsoid, vector $\mathbf{x}_a$ and $\mathbf{x}_b$ be the points on the surface of ellipsoids. $\mathbf{x}_a$, $\mathbf{x}_c$ and $\mathbf{X}_A$ are on the same line; and $\mathbf{x}_b$, $\mathbf{x}_c$ and $\mathbf{X}_B$ are on the same line.

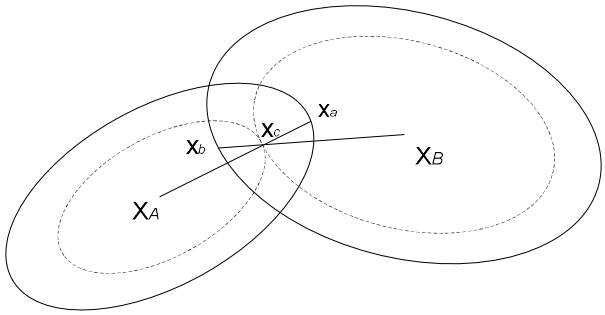



Figure 3.4.6 Two ellipsoid A and B. Two points $\mathbf{x}_a$ and $\mathbf{x}_b$ are on the surface of ellipsoid, and $\mathbf{x}_c$ is inside the ellipsoid. Point $\mathbf{x}_a$, $\mathbf{X}_A$ and $\mathbf{x}_c$ are on the same line; and the point $\mathbf{x}_a$, $\mathbf{X}_B$, and $\mathbf{x}_c$ are on the same line. The dotted line represents

the potential surface on which point $\mathbf{x}_c$ lies.

Similar to equation $\mathbf{x}_a = \mathbf{x}_A + \frac{\mathbf{x}_c - \mathbf{x}_A}{\sqrt{\mathbf{\Phi}_c}}$,

$$\begin{cases} \mathbf{x}_a = \mathbf{x}_A + \frac{\mathbf{x}_c - \mathbf{x}_A}{\sqrt{\mathbf{\Phi}_c}} \\ \mathbf{x}_b = \mathbf{x}_B + \frac{\mathbf{x}_c - \mathbf{x}_B}{\sqrt{\mathbf{\Phi}_c}} \end{cases} \tag{19}$$

Then I will do derivation so that all variables relative with radii of spheres in equation $W = \frac{2Ka^5}{5R_m^2}$, $a = \sqrt{R_m \delta}$, $\delta = R_A + R_B - \|\mathbf{x}_A - \mathbf{x}_B\|$ and $R_m = \frac{R_1 R_2}{R_1 + R_2}$ are replaced.

Firstly subtract the first equation by the second equation in $\begin{cases} \mathbf{x}_a = \mathbf{x}_A + \frac{\mathbf{x}_c - \mathbf{x}_A}{\sqrt{\mathbf{\Phi}_c}} \\ \mathbf{x}_b = \mathbf{x}_B + \frac{\mathbf{x}_c - \mathbf{x}_B}{\sqrt{\mathbf{\Phi}_c}} \end{cases}$,

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{x}_A - \mathbf{x}_B - \frac{1}{\sqrt{\mathbf{\Phi}_c}}(\mathbf{x}_A - \mathbf{x}_B) \tag{20}$$

Thus

$$\mathbf{x}_a - \mathbf{x}_b = \left(1 - \frac{1}{\sqrt{\mathbf{\Phi}_c}}\right)(\mathbf{x}_A - \mathbf{x}_B) \tag{21}$$

Denote

$$\mathbf{d} = \mathbf{x}_a - \mathbf{x}_b \tag{22}$$

Then equation $\mathbf{x}_a - \mathbf{x}_b = \left(1 - \frac{1}{\sqrt{\mathbf{\Phi}_c}}\right)(\mathbf{x}_A - \mathbf{x}_B)$ turns to

$$\mathbf{d} = \left(1 - \frac{1}{\sqrt{\mathbf{\Phi}}}\right)(\mathbf{x}_A - \mathbf{x}_B) \tag{23}$$

I consider $\mathbf{d}$ as the level of deformation of the two ellipsoids, thus $\delta$ in equation

$\delta = R_A + R_B - \left\| \mathbf{x}_A - \mathbf{x}_B \right\|$ can be replaced by $\mathbf{d}$.

$$\delta \Rightarrow d = \left( \frac{1}{\sqrt{\mathbf{\Phi}}} - 1 \right) \left\| \mathbf{x}_A - \mathbf{x}_B \right\| \tag{24}$$

Then I still need to replace $R_m$ in $R_m = \frac{R_1 R_2}{R_1 + R_2}$, in which $R_A R_B$ and $R_A + R_B$ need to be replaced separately.

It is known that the volume of sphere equals $V_{sphere} = \frac{4}{3}\pi r^3$, in which $r$ is the radius of the sphere. From ellipsoid volume formula $V_{ellipsoid} = \frac{4}{3}\pi a_1 a_2 a_3$. Then I consider the sphere as a special type of ellipsoid that 3 radii are the same. Thus the radii in term $R_A R_B$ can be replaced by the radii of the ellipsoid, as equation (89), (90) and (91) show

$$R_A \Rightarrow \left( a_1 a_2 a_3 \right)^{\frac{1}{3}} \tag{25}$$

$$R_B \Rightarrow \left( b_1 b_2 b_3 \right)^{\frac{1}{3}} \tag{26}$$

$$R_A R_B \Rightarrow \left( a_1 a_2 a_3 b_1 b_2 b_3 \right)^{\frac{1}{3}} \tag{27}$$

Note that the generalised Hertz formula should turn to normal Hertz formula with ellipsoids of 3 same radii. Thus I cannot replace term $R_A + R_B$ in the same way with $R_A R_B$. Denote

$$\left\| \mathbf{x}_A - \mathbf{x}_B \right\| \Rightarrow R \tag{28}$$

Then

$$\delta = \left( \frac{1}{\sqrt{\mathbf{\Phi}}} - 1 \right) R \tag{29}$$

$$R_A + R_B \Rightarrow \frac{R}{\sqrt{\mathbf{\Phi}}} \tag{30}$$

From $R_A R_B \Rightarrow \left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{3}}$ and $R_A + R_B \Rightarrow \frac{R}{\sqrt{\boldsymbol{\Phi}}}$,

$$R_m = \frac{\left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{3}}}{\frac{R}{\sqrt{\boldsymbol{\Phi}}}} \tag{31}$$

Substitute $\delta = \left(\frac{1}{\sqrt{\boldsymbol{\Phi}}} - 1\right) R$ and $R_m = \frac{\left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{3}}}{\frac{R}{\sqrt{\boldsymbol{\Phi}}}}$ to $a = \sqrt{R_m \delta}$,

$$\begin{aligned} a &= \sqrt{\frac{\left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{3}}}{\frac{R}{\sqrt{F\left(\mathbf{A},\mathbf{B},\mathbf{s}\right)}}} \left(\frac{R}{\sqrt{\boldsymbol{\Phi}}} - R\right)} \\ &= \left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{6}} \cdot \sqrt{\frac{\sqrt{\boldsymbol{\Phi}}}{R} \cdot R \frac{1-\sqrt{\boldsymbol{\Phi}}}{\sqrt{\boldsymbol{\Phi}}}} \\ &= \left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{6}} \sqrt{1-\sqrt{\boldsymbol{\Phi}}} \end{aligned} \tag{32}$$

Substitute $R_m = \frac{\left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{3}}}{\frac{R}{\sqrt{\boldsymbol{\Phi}}}}$ and $a = \sqrt{\frac{\left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{3}}}{\frac{R}{\sqrt{F\left(\mathbf{A},\mathbf{B},\mathbf{s}\right)}}} \left(\frac{R}{\sqrt{\boldsymbol{\Phi}}} - R\right)}$ $= \left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{6}} \cdot \sqrt{\frac{\sqrt{\boldsymbol{\Phi}}}{R} \cdot R \frac{1-\sqrt{\boldsymbol{\Phi}}}{\sqrt{\boldsymbol{\Phi}}}}$ $= \left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{6}} \sqrt{1-\sqrt{\boldsymbol{\Phi}}}$ to

$W = \frac{2Ka^5}{5R_m^2}$,

$$W = \frac{2Ka^5}{5R_m^2} = \frac{2K\left[\left(a_1 a_2 a_3\right)^{\frac{1}{3}} + \left(b_1 b_2 b_3\right)^{\frac{1}{3}}\right]^2 \left(a_1 a_2 a_3 b_1 b_2 b_3\right)^{\frac{1}{6}}}{5} \cdot \left[1-\sqrt{\boldsymbol{\Phi}}\right]^{\frac{5}{2}} \tag{33}$$

As discussed in Section 2.4.4, the value of force equals the partial differential of

energy against the displacement. So that

$$\mathbf{F}^{A} = -\frac{\partial W}{\partial \mathbf{s}} = -\frac{\partial W}{\partial \Phi}\frac{\partial \Phi}{\partial \mathbf{s}} \tag{34}$$

It is known that

$$\frac{\partial W}{\partial \Phi} = \frac{2K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}} + \left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2 \left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}}{5} \cdot \frac{5}{2}\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}} \cdot \left(-\frac{1}{2}\right)\Phi^{-\frac{1}{2}} \tag{35}$$

And

$$\frac{\partial \Phi}{\partial \mathbf{s}} = 2\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s} \tag{36}$$

Substitute equation

$$\frac{\partial W}{\partial \Phi} = \frac{2K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}} + \left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2 \left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}}{5} \cdot \frac{5}{2}\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}} \cdot \left(-\frac{1}{2}\right)\Phi^{-\frac{1}{2}}$$

and

$\frac{\partial \Phi}{\partial \mathbf{s}} = 2\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}$ to $\mathbf{F}^{A} = -\frac{\partial W}{\partial \mathbf{s}} = -\frac{\partial W}{\partial \Phi}\frac{\partial \Phi}{\partial \mathbf{s}}$,

$$\mathbf{F}^{con} = -\frac{\partial W}{\partial \mathbf{s}} = -\frac{\partial W}{\partial \Phi}\frac{\partial \Phi}{\partial \mathbf{s}}$$

$$= -\frac{2K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}} + \left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2 \left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}}{5} \cdot \frac{5}{2}\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}} \cdot \left(-\frac{1}{2}\right)\Phi^{-\frac{1}{2}} \cdot \left(2\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right)$$

$$= K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}} + \left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2 \left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}} \cdot \left(1-\sqrt{\Phi}\right)^{\frac{3}{2}} \cdot \Phi^{-\frac{1}{2}} \cdot \left(\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right) \tag{37}$$

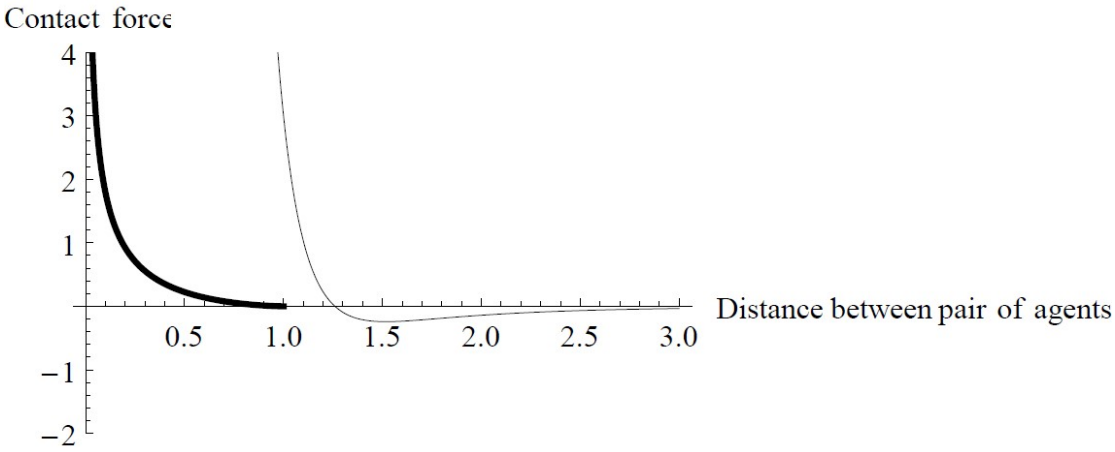



Figure 3.4.7 Size of contact force from Lenard potential (thin black line) and Hertz formula (thick black line)

As Figure 3.4.7 shows, the contact force calculated from Hertz model increases much slower than contact force calculated from Lenard potential, which means Hertz model can be used to present much more elastic agents, while Lenard potential describes rigid agents. Therefore to describe the interaction between pair of cells, Hertz model is more suitable. Equation

$$\mathbf{F}^{con} = -\frac{\partial W}{\partial \mathbf{s}} = -\frac{\partial W}{\partial \Phi}\frac{\partial \Phi}{\partial \mathbf{s}}$$

$$= -\frac{2K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}} + \left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2 \left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}}{5} \cdot \frac{5}{2}\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}} \cdot \left(-\frac{1}{2}\right)\Phi^{-\frac{1}{2}} \cdot \left(2\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right)$$

$$= K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}} + \left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2 \left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}} \cdot \left(1-\sqrt{\Phi}\right)^{\frac{3}{2}} \cdot \Phi^{-\frac{1}{2}} \cdot \left(\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right)$$

is the contact force derived from generalised Hertz formula. The calculation is complex. Like the adhesion force, the contact force should be calculated for each pair of agents. The sum of adhesion and contact force will be used to calculate agent velocity in Section 3.6.

## 3.5 Contact torque and adhesion torque

The contact and adhesion torques are the result of my assumption that the shape of agent is ellipsoid. The torques are used to describe the rotation of agents along an arbitrary axis, as Figure 3.5.1 shows. This axis may be coincident with one of the three semi-axes of the ellipsoid but it also may not; also generally speaking the axis may not pass the centre of the ellipsoid, which makes the rotation a complex problem. From the torque the angular velocity of the agent can be calculated, which represent the speed that agents rotates along the axis.

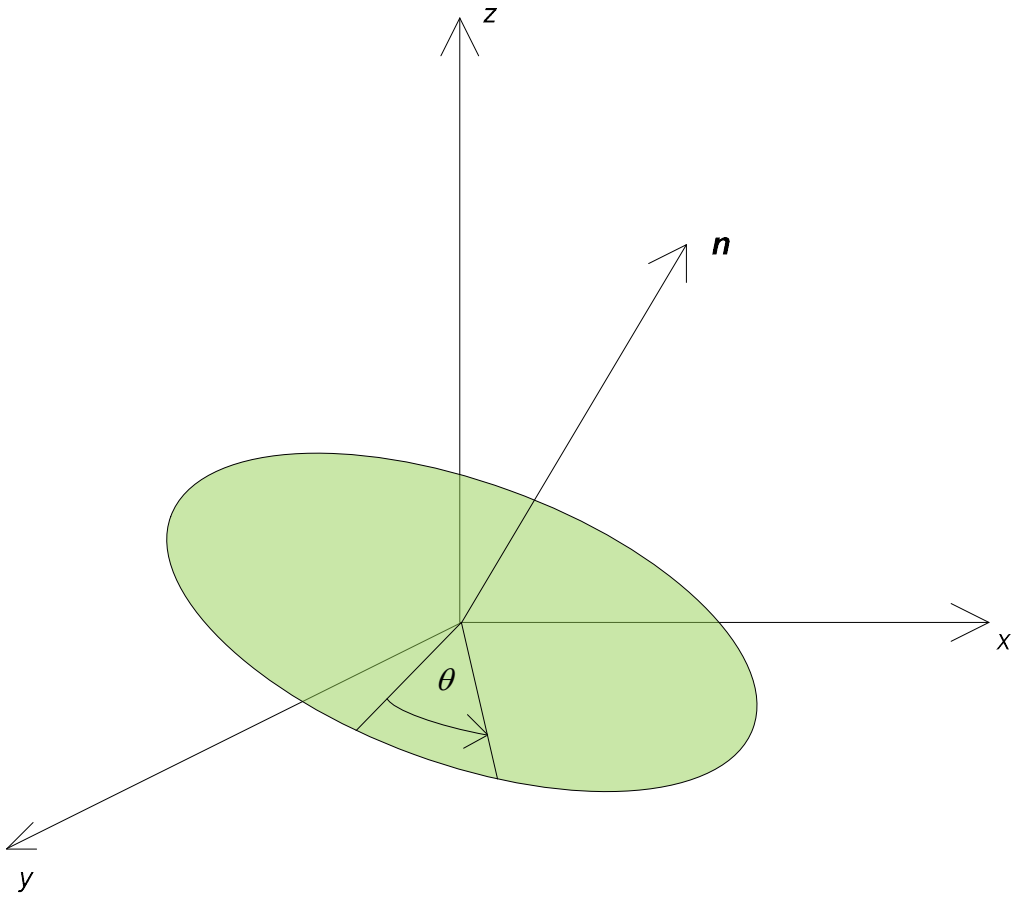



Figure 3.5.1 Rotation by angle $\theta$ along axis $\mathbf{n}$

I represent a rotation of angle $\theta$ along arbitrary axis $\mathbf{n}$ as $\mathbf{R}_{\mathbf{n}}^{\theta}$, an arbitrary vector

as $\mathbf{r}$, after the rotation $\mathbf{r}$ turns to $\mathbf{r}'$. Then $\mathbf{r}'$ and $\mathbf{r}$ satisfy

$$\mathbf{r}' = \mathbf{R}_{\mathbf{n}}^{\theta}\mathbf{r} \tag{1}$$

$\mathbf{R}_{\mathbf{n}}^{\theta}$ is a matrix as follows

$$\mathbf{R} = \begin{pmatrix} n_1^2(1-\cos\theta)+\cos\theta & n_1n_2(1-\cos\theta)-n_3\sin\theta & n_1n_3(1-\cos\theta)+n_2\sin\theta \\ n_1n_2(1-\cos\theta)+n_3\sin\theta & n_2^2(1-\cos\theta)+\cos\theta & n_2n_3(1-\cos\theta)-n_1\sin\theta \\ n_1n_3(1-\cos\theta)-n_2\sin\theta & n_2n_3(1-\cos\theta)+n_1\sin\theta & n_3^2(1-\cos\theta)+\cos\theta \end{pmatrix} \tag{2}$$

In which $\theta$ is the rotation angle; $n_1$, $n_2$, and $n_3$ are coeffients.

Then $\mathbf{R}_{\mathbf{n}}^{\theta}$ can also be expressed in component form, which is a compact form of matrix form. The component form is useful in differential calculation.

$$R_{kl} = \delta_{kl}\cos\theta - \varepsilon_{klm}\sin\theta n_m + (1-\cos\theta)n_k n_l \tag{3}$$

In which $\varepsilon_{klm}$ is a three-dimensional Levi-Civita symbol, and value of $\varepsilon_{klm}$ satisfies

$$\varepsilon_{klm} = \begin{cases} +1 \text{ if } (k,l,m) = (1,2,3) \text{ or } (2,3,1) \text{ or } (3,1,2) \\ -1 \text{ if } (k,l,m) = (1,3,2) \text{ or } (3,2,1) \text{ or } (2,1,3) \\ 0 \text{ otherwise} \end{cases} \tag{4}$$

Similar to calculation of the forces $\mathbf{F} = \dfrac{\partial W}{\partial(\mathbf{A}-\mathbf{B})}$, the calculation of torque can be written as $\boldsymbol{\tau} = \dfrac{\partial W}{\partial \mathbf{u}}$, in which $W$ is the energy and $\mathbf{u}$ is the direction of the axis. First, I derive an expression for the rotation along arbitrary axis. Let $\mathbf{R}_{\mathbf{n}}^{\delta\beta}$ be an infinitesimal rotation angle $\delta\beta$ around $\mathbf{n}$ axis, here $\mathbf{n}$ is a normal direction along rotation axis, denote $\delta\boldsymbol{\beta} = \mathbf{n}\delta\beta$, in which $\delta\boldsymbol{\beta}$ is a vector representing rotation of $\delta\beta$

along axis $\mathbf{n}$. Then similar to equation $R_{kl} = \delta_{kl}\cos\theta - \varepsilon_{klm}\sin\theta n_m + (1-\cos\theta)n_k n_l$, $\mathbf{R}_{\mathbf{n}}^{\delta\beta}$ can be expressed in component form as

$$R_{kl}(\delta\boldsymbol{\beta}) = \delta_{kl} - \varepsilon_{klm}\delta\beta_m = \delta_{kl} - \delta\beta\varepsilon_{klm}n_m \tag{5}$$

To get a expression of rotation $\delta\beta$, multiply $\varepsilon_{kln}$ to both end of $R_{kl}(\delta\boldsymbol{\beta}) = \delta_{kl} - \varepsilon_{klm}\delta\beta_m = \delta_{kl} - \delta\beta\varepsilon_{klm}n_m$, and get $R_{kl} = 0 - 2\delta\beta_n$ here the suffix 'n' is the result of calculation of Levi-Civita symbol. The inverse relationship of Equation $R_{kl}(\delta\boldsymbol{\beta}) = \delta_{kl} - \varepsilon_{klm}\delta\beta_m = \delta_{kl} - \delta\beta\varepsilon_{klm}n_m$ is

$$\delta\beta_m = \delta\beta n_m = -\frac{1}{2}\varepsilon_{klm}R_{kl}(\delta\boldsymbol{\beta}) \tag{6}$$

Then rotated arbitrary vector $\mathbf{r}$ becomes new vector $\mathbf{r}'$, their relationship can be written as

$$\mathbf{r}' = \mathbf{r}(\boldsymbol{\beta} + \delta\boldsymbol{\beta}) = \mathbf{R}_{\mathbf{n}}^{\delta\beta}\cdot\mathbf{r} \tag{7}$$

Thus the change of direction of vector $\mathbf{r}$ is

$$\delta\mathbf{r} = \mathbf{r}' - \mathbf{r} = \mathbf{r}(\boldsymbol{\beta} + \delta\boldsymbol{\beta}) - \mathbf{r}(\boldsymbol{\beta}) = \left[\mathbf{R}(\delta\boldsymbol{\beta}) - \mathbf{I}\right]\cdot\mathbf{r}(\boldsymbol{\beta}) \tag{8}$$

Write $\delta\mathbf{r} = \mathbf{r}' - \mathbf{r} = \mathbf{r}(\boldsymbol{\beta} + \delta\boldsymbol{\beta}) - \mathbf{r}(\boldsymbol{\beta}) = \left[\mathbf{R}(\delta\boldsymbol{\beta}) - \mathbf{I}\right]\cdot\mathbf{r}(\boldsymbol{\beta})$ in component form

$$\delta r_k = r_k(\boldsymbol{\beta} + \delta\boldsymbol{\beta}) - r_k(\boldsymbol{\beta}) = -\delta\beta\varepsilon_{klm}n_m r_l \tag{9}$$

Equation $\delta r_k = r_k(\boldsymbol{\beta} + \delta\boldsymbol{\beta}) - r_k(\boldsymbol{\beta}) = -\delta\beta\varepsilon_{klm}n_m r_l$ is relationship of change of vector $\mathbf{r}$ and the rotation $\delta\beta$. Then $\delta r_k = r_k(\boldsymbol{\beta} + \delta\boldsymbol{\beta}) - r_k(\boldsymbol{\beta}) = -\delta\beta\varepsilon_{klm}n_m r_l$ can be written as

$$\frac{\partial r_k}{\partial\beta} = -\varepsilon_{klm}n_m r_l = \varepsilon_{klm}n_l r_m \tag{10}$$

Which means

$$\frac{\partial \mathbf{r}(\boldsymbol{\beta})}{\partial \beta} = \mathbf{n} \times \mathbf{r} \tag{11}$$

Expand vector $\mathbf{r}$ to the direction matrix of ellipsoid $\mathbf{u}_l$, the relationship between change of direction of an arbitrary ellipsoid rotated and rotation angle $\delta\beta$ is known:

$$\frac{\partial \mathbf{u}_l}{\partial \beta} = \mathbf{n} \times \mathbf{u}_l \quad (l = 1,2,3) \tag{12}$$

In which $\mathbf{n}$ is rotation axis.

Or in component form

$$\frac{\partial u_{kl}}{\partial \beta} = \varepsilon_{kpq} n_p u_{ql} \tag{13}$$

Similar to equation (2.3) in (Allen and Germano, 2002), the torque of ellipsoid A in pair of ellipsoids satisfies

$$\mathbf{n} \cdot \boldsymbol{\tau}^A = n_p \tau_p^A = -\frac{\partial W}{\partial \beta} = -\frac{\partial W}{\partial u_{kl}} \frac{\partial u_{kl}}{\partial \beta} = -\frac{\partial W}{\partial u_{kl}} \varepsilon_{kpq} n_p u_{ql} \tag{14}$$

In which $W$ is the energy, $u_{kl}$ is the direction of the ellipsoid. Then

$$\tau_p^A = -\frac{\partial W}{\partial u_{kl}} \varepsilon_{kpq} u_{ql} \tag{15}$$

Similarly the other ellipsoid in pair of ellipsoids has torque as

$$\tau_p^B = -\frac{\partial W}{\partial v_{kl}} \varepsilon_{kpq} v_{ql} \tag{16}$$

Note that $W(\mathbf{A},\mathbf{B},\mathbf{s}) = W\left[\Phi(\mathbf{A},\mathbf{B},\mathbf{s})\right]$, which means the energy is function of potential $\Phi$. It is not convenient to calculate $\frac{\partial W}{\partial \mathbf{u}}$, however according to characteristic of partial differential equation, $\frac{\partial W}{\partial \mathbf{u}} = \frac{\partial W}{\partial \mathbf{\Phi}} \cdot \frac{\partial \mathbf{\Phi}}{\partial \mathbf{u}}$. Thus

$$\tau_p^A = -\frac{\partial W}{\partial \Phi}\frac{\partial \Phi}{\partial u_{kl}}\varepsilon_{kpq}u_{ql} \tag{17}$$

$$\tau_p^B = -\frac{\partial W}{\partial \Phi}\frac{\partial \Phi}{\partial v_{kl}}\varepsilon_{kpq}v_{ql} \tag{18}$$

From Equation $\Phi(\mathbf{A},\mathbf{B},\mathbf{s}) = F(\lambda_0) = \lambda_0(1-\lambda_0)\mathbf{s}^T\mathbf{G}^{-1}(\lambda_0)\mathbf{s}$, and note that $\mathbf{G}^{-1}\mathbf{G} = \mathbf{I}$

$$\frac{\partial \Phi}{\partial \mathbf{U}} = \lambda_0(1-\lambda_0)\mathbf{s}^T\frac{\partial \mathbf{G}^{-1}(\lambda_0)}{\partial \mathbf{U}}\mathbf{s} = -\lambda_0(1-\lambda_0)\mathbf{s}^T\mathbf{G}^{-1}(\lambda_0)\frac{\partial \mathbf{G}(\lambda_0)}{\partial \mathbf{U}}\mathbf{G}^{-1}(\lambda_0)\mathbf{s} \tag{19}$$

In which U is the direction matrix of ellipsoid. Then it is needed to derive $\frac{\partial \mathbf{G}(\lambda_0)}{\partial \mathbf{U}}$.

Using $\mathbf{G}(\lambda) = \lambda\mathbf{B}^{-1} + (1-\lambda)\mathbf{A}^{-1}$,

$$\frac{\partial \Phi}{\partial \mathbf{U}} = -\lambda_0(1-\lambda_0)^2\mathbf{s}^T\mathbf{G}^{-1}(\lambda_0)\frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{U}}\mathbf{G}^{-1}(\lambda_0)\mathbf{s} \tag{20}$$

Or in component form

$$\frac{\partial \Phi}{\partial u_{kl}} = -\lambda_0(1-\lambda_0)^2 s_i G_{im}^{-1}(\lambda_0)\frac{\partial A_{mn}^{-1}}{\partial u_{kl}}G_{nj}^{-1}(\lambda_0)s_j \tag{21}$$

From characteristic of **A** it is known

$$\mathbf{A}^{-1} = \mathbf{U}\overline{\mathbf{A}}^{-1}\mathbf{U}^T = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix}\begin{pmatrix} a_1^2 & 0 & 0 \\ 0 & a_2^2 & 0 \\ 0 & 0 & a_3^2 \end{pmatrix}\begin{pmatrix} u_{11} & u_{21} & u_{31} \\ u_{12} & u_{22} & u_{32} \\ u_{13} & u_{23} & u_{33} \end{pmatrix} \tag{22}$$

i.e.

$$A_{mn}^{-1} = u_{mr}\overline{A}_{rs}^{-1}u_{ns} \tag{23}$$

$$\frac{\partial A_{mn}^{-1}}{\partial u_{kl}} = \delta_{km}\delta_{lr}\overline{A}_{rs}^{-1}u_{ns} + u_{mr}\overline{A}_{rs}^{-1}\delta_{kn}\delta_{ls} \tag{24}$$

Substituting equation $\frac{\partial A_{mn}^{-1}}{\partial u_{kl}} = \delta_{km}\delta_{lr}\overline{A}_{rs}^{-1}u_{ns} + u_{mr}\overline{A}_{rs}^{-1}\delta_{kn}\delta_{ls}$ into Equation

$$\frac{\partial \Phi}{\partial u_{kl}} = -\lambda_0\left(1-\lambda_0\right)^2 s_i G_{im}^{-1}\left(\lambda_0\right)\frac{\partial A_{mn}^{-1}}{\partial u_{kl}}G_{nj}^{-1}\left(\lambda_0\right)s_j$$

$$\frac{\partial \Phi}{\partial u_{kl}} = -\lambda_0\left(1-\lambda_0\right)^2\left\{\begin{array}{l}\left[s_i G_{ik}^{-1}\left(\lambda_0\right)\right]\left[\overline{A}_{ls}^{-1}u_{ns}G_{nj}^{-1}\left(\lambda_0\right)s_j\right]+\\ \left[s_i G_{im}^{-1}\left(\lambda_0\right)u_{mr}\overline{A}_{rl}^{-1}\right]\left[G_{kj}^{-1}\left(\lambda_0\right)s_j\right]\end{array}\right\} \tag{25}$$

Note that both $\mathbf{G}^{-1}$ and $\overline{\mathbf{A}}$ are symmetric matrix.

$$\frac{\partial \Phi}{\partial u_{kl}} = -2\lambda_0\left(1-\lambda_0\right)^2\left[G_{ki}^{-1}\left(\lambda_0\right)s_i\right]\left[\overline{A}_{lr}^{-1}u_{nr}G_{nj}^{-1}\left(\lambda_0\right)s_j\right] \tag{26}$$

Or in block form

$$\frac{\partial \Phi}{\partial \mathbf{U}} = -2\lambda_0\left(1-\lambda_0\right)^2\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\otimes\left[\overline{\mathbf{A}}^{-1}\mathbf{U}^T\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right] \tag{27}$$

Substituting equation $\frac{\partial \Phi}{\partial u_{kl}} = -2\lambda_0\left(1-\lambda_0\right)^2\left[G_{ki}^{-1}\left(\lambda_0\right)s_i\right]\left[\overline{A}_{lr}^{-1}u_{nr}G_{nj}^{-1}\left(\lambda_0\right)s_j\right]$ into equation $\tau_p^A = -\frac{\partial W}{\partial u_{kl}}\varepsilon_{kpq}u_{ql}$

$$\tau_p^A = 2\lambda_0\left(1-\lambda_0\right)^2\frac{\partial W}{\partial \Phi}\varepsilon_{pqk}\left[A_{qn}^{-1}G_{nj}^{-1}\left(\lambda_0\right)s_j\right]\left[G_{ki}^{-1}\left(\lambda_0\right)s_i\right] \tag{28}$$

Similarly

$$\tau_p^B = 2\lambda_0^2\left(1-\lambda_0\right)\frac{\partial W}{\partial \Phi}\varepsilon_{pqk}\left[B_{qn}^{-1}G_{nj}^{-1}\left(\lambda_0\right)s_j\right]\left[G_{ki}^{-1}\left(\lambda_0\right)s_i\right] \tag{29}$$

Or in block form, the torques of two non-identical ellipsoids **A** and **B** are

$$\boldsymbol{\tau}^{A} = 2\lambda_0\left(1-\lambda_0\right)^2 \frac{\partial W}{\partial \Phi}\left[\bar{\mathbf{A}}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right] \tag{30}$$

$$\boldsymbol{\tau}^{B} = 2\lambda_0^2\left(1-\lambda_0\right) \frac{\partial W}{\partial \Phi}\left[\bar{\mathbf{B}}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right] \tag{31}$$

Where $W$ represents the energy and is the function of potential $\Phi$ ( $\Phi$ is the function of geometry information of ellipsoid A and B, and distance vector $\mathbf{s}$ between the two ellipsoids)

$$W\left(\mathbf{A},\mathbf{B},\mathbf{s}\right) = W\left[\boldsymbol{\Phi}\left(\mathbf{A},\mathbf{B},\mathbf{s}\right)\right] \tag{32}$$

To calculate contact torque, the energy equation $W = \dfrac{2Ka^5}{5R_m^2} = \dfrac{2K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}}+\left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2\left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}}{5}\cdot\left[1-\sqrt{\boldsymbol{\Phi}}\right]^{\frac{5}{2}}$ is substituted to equation $\boldsymbol{\tau}^{A} = 2\lambda_0\left(1-\lambda_0\right)^2 \dfrac{\partial W}{\partial \Phi}\left[\bar{\mathbf{A}}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]$

$$\begin{aligned}
\boldsymbol{\tau}^{Con} &= 2\lambda_0\left(1-\lambda_0\right)^2 \frac{\partial W}{\partial \Phi}\left[\mathbf{A}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right] \\
&= 2\lambda_0\left(1-\lambda_0\right)^2\left\{\frac{2K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}}+\left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2\left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}}{5}\cdot\frac{5}{2}\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}}\cdot\left(-\frac{1}{2}\right)\Phi^{-\frac{1}{2}}\right\} \\
&\cdot\left[\mathbf{A}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right] \\
&= -\lambda_0\left(1-\lambda_0\right)^2 K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}}+\left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2\left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}\cdot\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}}\cdot\Phi^{-\frac{1}{2}} \\
&\cdot\left[\mathbf{A}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]
\end{aligned} \tag{33}$$

Similarly substitute adhesion potential $\frac{\partial W^{Ad}}{\partial \Phi}=\frac{\partial\left(\varepsilon\sqrt{\Phi}\right)}{\partial \Phi}=\frac{1}{2}\varepsilon\Phi^{-\frac{1}{2}}$ to $\boldsymbol{\tau}^{A}=2\lambda_0\left(1-\lambda_0\right)^2\frac{\partial W}{\partial \Phi}\left[\bar{\mathbf{A}}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]$, and adhesion torque can be written as

$$\boldsymbol{\tau}^{Ad}=\varepsilon\lambda_0\left(1-\lambda_0\right)^2\Phi^{-\frac{1}{2}}\left[\mathbf{A}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right] \tag{34}$$

In which $\times$ is cross produce. As Figure 3.5.2 shows, the contact and adhesion torque drive agents to rotate until agents are parallel.

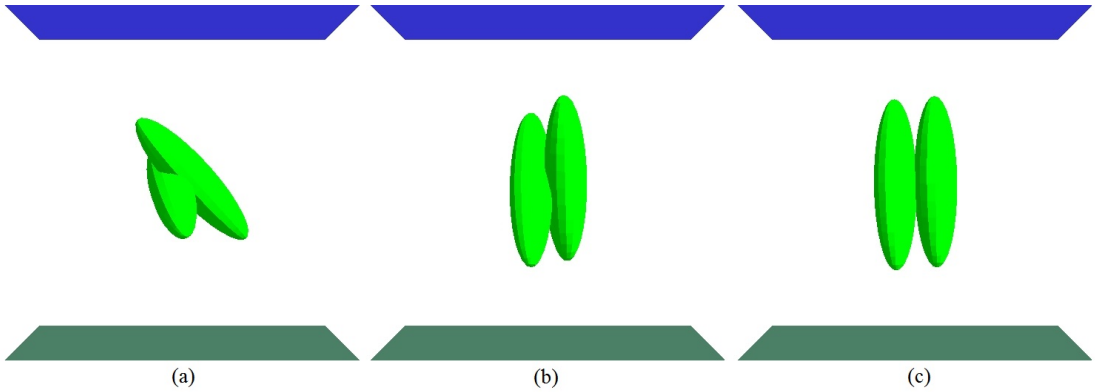



Figure 3.5.2 From (a) to (c), two agents (green ellipslidal particle) under effect of contact and adhesion torque rotate until they are parallel.

## 3.6 Velocity and angular velocity

According to Section 2.3, cells can be considered as particles moving in fluid (Drasdo and Hohme, 2005). Thus to calculate velocity and angular velocity, the Stokes resistance rule may be applied. The Stokes resistance force and torque are in the following form (Derjaguin *et al*. 1975).

$$\mathbf{F}^{s}=-\mu\mathbf{K}^{t}\cdot\left(\mathbf{u}-\mathbf{u}_{f}\right) \tag{1}$$

$$\mathbf{T}^{s}=-\mu\mathbf{K}^{r}\cdot\left(\boldsymbol{\omega}-\boldsymbol{\omega}_{f}\right) \tag{2}$$

In which $\mathbf{u}$ and $\boldsymbol{\omega}$ are the velocity and angular velocity of an ellipsoid respectively, $\mathbf{u}_f$ and $\mathbf{w}_f$ are the velocity and angular velocity of fluid respectively, and $\mu$ is the viscosity coefficiency. As I assume the fluid in which cells grow and migrate is still, i.e. the fluid has neither velocity nor angular velocity. Thus $\mathbf{u}_f$ and $\mathbf{w}_f$ are zero vectors, and equation $\mathbf{F}^s = -\mu\mathbf{K}^t \cdot \left(\mathbf{u} - \mathbf{u}_f\right)$ and $\mathbf{T}^s = -\mu\mathbf{K}^r \cdot \left(\boldsymbol{\omega} - \boldsymbol{\omega}_f\right)$ turn to

$$\mathbf{F}^s = -\mu\mathbf{K}^t\mathbf{u} \tag{3}$$

$$\mathbf{T}^s = -\mu\mathbf{K}^r\boldsymbol{\omega} \tag{4}$$

$\mathbf{K}^t$ and $\mathbf{K}^r$ are Stokes resistance (Brenner, H. 1963) constants that satisfy

$$\mathbf{K}^t = 16\pi\left(\frac{1}{\chi + a_1^2\alpha_1}\mathbf{u}_1 \otimes \mathbf{u}_1 + \frac{1}{\chi + a_2^2\alpha_2}\mathbf{u}_2 \otimes \mathbf{u}_2 + \frac{1}{\chi + a_3^2\alpha_3}\mathbf{u}_3 \otimes \mathbf{u}_3\right) \tag{5}$$

$$\mathbf{K}^r = 16\pi\left(\frac{a_2^2 + a_3^2}{a_2^2\alpha_2 + a_3^2\alpha_3}\mathbf{u}_1 \otimes \mathbf{u}_1 + \frac{a_1^2 + a_3^2}{a_1^2\alpha_1 + a_3^2\alpha_3}\mathbf{u}_2 \otimes \mathbf{u}_2 + \frac{a_2^2 + a_1^2}{a_2^2\alpha_2 + a_1^2\alpha_1}\mathbf{u}_3 \otimes \mathbf{u}_3\right) \tag{6}$$

In which $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ are the directions of each semi-axis of an ellipsoid cell, and $a_1$, $a_2$ and $a_3$ are the lengths of each semi-axis. The $\alpha_1$, $\alpha_2$ and $\alpha_3$ can be evaluated as follows, in which $a_1$, $a_2$ and $a_3$ are three semi-axis of ellipsoid and satisfy $a_1 < a_2 < a_3$. Done

$$\alpha_k = \int_0^\infty \frac{d\lambda}{\left(a_k^2 + \lambda\right)\Delta\lambda} \quad (k = 1,2,3) \tag{7}$$

$$\chi = \int_0^\infty \frac{d\lambda}{\Delta\lambda} \quad (k = 1,2,3) \tag{8}$$

$$\Delta\lambda = \sqrt{\left(a_1^2 + \lambda\right)\left(a_2^2 + \lambda\right)\left(a_3^2 + \lambda\right)} \tag{9}$$

For any ellipsoid in fluid, the sum of contact force and adhesion force should balance

the Stokes resistance $\mathbf{F}^s$ and therefore satisfy

$$\sum \mathbf{F} + \mathbf{F}^s = 0 \tag{10}$$

From equation $\mathbf{F}^s = -\mu \mathbf{K}^t \mathbf{u}$ and $\sum \mathbf{F} + \mathbf{F}^s = 0$,

$$\mathbf{F}^s = -\sum \mathbf{F} = -\mu \mathbf{K}^t \mathbf{u} \tag{11}$$

So that

$$\mathbf{u} = \frac{1}{\mu}\left[\mathbf{K}^t\right]^{-1} \sum \mathbf{F} \tag{12}$$

Similar to $\mathbf{F}^s = -\mu \mathbf{K}^t \mathbf{u}$, the sum of contact torque and adhesion torque should balance the Stokes resistance $\mathbf{T}^s$

$$\sum \mathbf{T} + \mathbf{T}^s = 0 \tag{13}$$

From $\mathbf{T}^s = -\mu \mathbf{K}^r \boldsymbol{\omega}$ and $\sum \mathbf{T} + \mathbf{T}^s = 0$,

$$\boldsymbol{\omega} = \frac{1}{\mu}\left[\mathbf{K}^r\right]^{-1} \sum \mathbf{T} \tag{14}$$

In this way I get the velocity and angular velocity of the cell from the total force/torque working on it.

## 3.7 Cell position and direction update

Given the velocity and angular velocity of a cell and an initial position $\mathbf{P}_0$ and calculated velocity $\mathbf{u}$, the next position after time $\Delta t$ can be calculated as

$\mathbf{P}_1 = \mathbf{P}_0 + \mathbf{u} \cdot \Delta t$ . Likewise, given the initial direction of an ellipsoid $\mathbf{R}_0$ and the calculated angular velocity $\boldsymbol{\omega}$, the next direction after time $\Delta t$ can be calculated as $\mathbf{R}_1 = (\boldsymbol{\omega} \cdot \Delta t) \cdot \mathbf{R}_0$.

## 3.8 Cell-substrate interaction

Consider the situation of two mirror symmetry identical agents; it can be seen that the interactions are also mirrored: the size of forces and torques are the same, and their directions are mirrored in symmetry, as Figure 3.8.2(a) shows. Further, if I consider the planar substrate as the 'mirror' and the interaction between an agent and the substrate plane as the interaction between two mirror symmetry identical agents (shown in Figure 3.8.2(b)), the interactions shown in Figure 3.8.2(a) and Figure 3.8.2(b) become exchangable. In this thesis the interaction in Figure 3.8.2(b) is simplified and considered as the interaction between an agent and its mirror image.

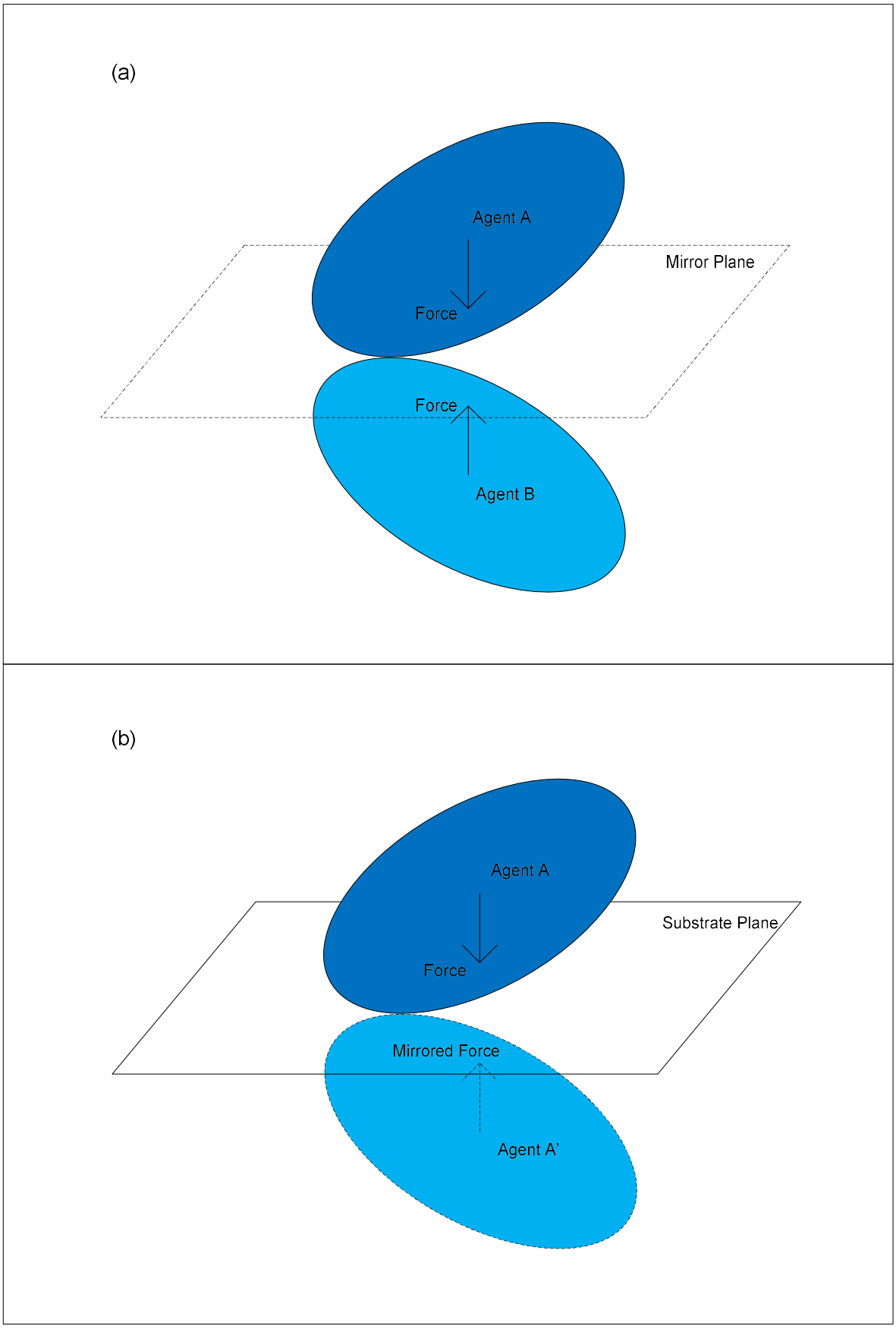



Figure 3.8.2 (a) pair of identical agents A and B, can be considered in mirror

symmetry by a mirror plane; (b) an agent A interacts with the substrate plane represented as the mirror image of the agent A'.

There are traditional equations to describe the interaction between an elastic ellipsoid and a substrate (Brenner and Gajdos, 1980). However in this study the potential between the ellipsoidal particle and substrate is presented in similar shape of equation $V(r)=4\varepsilon\left[\left(\frac{\sigma^2}{r^2}\right)^6-\left(\frac{\sigma^2}{r^2}\right)^3\right]$, which makes it suitable to describe a rigid particle, not an elastic particle like cell. The simplified method may not be accurate to describe the cell-substrate interaction, but by simplifying the cell-substrate interaction as the interaction between the agent and its mirror image, I gain a convenient consistency (same level of simplify) between inter-agent and agent-environment interactions. The Hertz formula can be used to calculate contact force between an agent and a substrate plane (H. R. Hertz, 1882). There are potential problems with the adhesion force calculation: the properties of the substrate are not introduced to the model, thus it is hard to calibrate the adhesion force in my model with experimental measurements. It is also hard for the model to simulate different types of agent- substrate adhesion forces or test the response the react of agent to different types of substrates. These problems lie outside of the scope of this thesis.

To simplify the cell-substrate calculation, cell-substrate interaction is considered as the interaction between a cell and its mirror image by the substrate plane. As Figure 3.8.2(b) shows, $\mathbf{A}'$ is not a real cell, but the mirror image of cell $\mathbf{A}$. So all the force and torque on cell A between it and the substrate can be calculated between A and $\mathbf{A}'$.

Denote the ellipsoid as ellipsoid $\mathbf{A}$ and its mirror image ellipsoid $\mathbf{A}'$,

$$\mathbf{A}=\begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix}\begin{pmatrix} \frac{1}{a_1^2} & 0 & 0 \\ 0 & \frac{1}{a_2^2} & 0 \\ 0 & 0 & \frac{1}{a_3^2} \end{pmatrix}\begin{pmatrix} u_{11} & u_{21} & u_{31} \\ u_{12} & u_{22} & u_{32} \\ u_{13} & u_{23} & u_{33} \end{pmatrix} \tag{1}$$

As $\mathbf{A}'$ is mirror image of $\mathbf{A}$, they are identical in size and shape, but the direction of semi-axis of $\mathbf{A}'$ has an opposite z-component.

$$\mathbf{A}'=\begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ -u_{31} & -u_{32} & -u_{33} \end{pmatrix}\begin{pmatrix} \frac{1}{a_1^2} & 0 & 0 \\ 0 & \frac{1}{a_2^2} & 0 \\ 0 & 0 & \frac{1}{a_3^2} \end{pmatrix}\begin{pmatrix} u_{11} & u_{21} & -u_{31} \\ u_{12} & u_{22} & -u_{32} \\ u_{13} & u_{23} & -u_{33} \end{pmatrix} \tag{2}$$

And $\mathbf{r}_A\left(x_A \quad y_A \quad z_A\right)$, $\mathbf{r}_B\left(x_A \quad y_A \quad -z_A\right)$, so

$$\mathbf{F}^{slab}=\left[\frac{5}{2}K\cdot\left(a_1a_2a_3b_1b_2b_3\right)^{\frac{5}{6}}\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}}\cdot\Phi^{-\frac{1}{2}}\right]\cdot\left[\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\right] \tag{3}$$

And

$$\begin{aligned}\tau_A^{slab}&=-\left[K\cdot\left(a_1a_2a_3b_1b_2b_3\right)^{\frac{5}{6}}\cdot\frac{5}{2}\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}}\cdot\left(-\frac{1}{2}\right)\Phi^{-\frac{1}{2}}\right]\cdot\\ &\left(-2\right)\lambda_0\left(1-\lambda_0\right)^2\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\otimes\left[\mathbf{A}^{-1}\mathbf{U}^T\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\\ &=\frac{5}{2}K\cdot\left(a_1a_2a_3b_1b_2b_3\right)^{\frac{5}{6}}\cdot\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}}\cdot\Phi^{-\frac{1}{2}}\cdot\lambda_0\left(1-\lambda_0\right)^2\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\otimes\left[\mathbf{A}^{-1}\mathbf{U}^T\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\end{aligned} \tag{4}$$

However the size of force and torque between agent and substrate equals that between

an agent and another agent. In the case where one agent interacts with a small number of other agents, the agent-agent force may become too large and pull the single agent away from the substrate, which does not match the *in vitro* experiment. To solve this problem I assume the agent-substrate interaction is 5 times of agent-agent interaction. This value, determined by model parameter exploration, allows agents to move across the substrate surface without detaching from the substrate itself.

## 3.9 Dimension analysis

Prior to implementation of the model for use in simulation, there are several parameters, such as the constant in Hertz formula, that need a real-world value. More fundamentally, it is important to determine if, in all equations, no matter whether it is for force, torque or velocity, all the parameters have their own physical interpretation which can be linked to its dimension. Thus it is needed to ensure every physical property in the model has a correct dimension. Dimension analysis (Bhaskar and Nigam, 1990), (Grigory Isaakovich Barenblatt, 1996) is a method that enables exactly this. For full details please refer to Appendix B Step 1, in which dimension of all physical entities are tested. The list of physical entities and their dimension is shown in Table 3.9.1.

| Name of physical entity | Symbol | Dimension |
|---|---|---|
| Contact potential variable | $\lambda_0$ | 1 |
| Distance between pair of agents | $\mathbf{s}$ | $\Delta x$ |
| Shape of ellipsoidal agent | $\mathbf{A}$ | $\frac{1}{\Delta x^2}$ |
| Semi-axes of ellipsoidal agent | $a_k \left(k = 1,2,3\right)$ | $\Delta x$ |

| | | |
|---|---|---|
| Intermediate matrix of contact potential | $\mathbf{G}(\lambda)$ | $\Delta x^2$ |
| Contact potential | $\mathbf{\Phi}(\mathbf{A},\mathbf{B},\lambda)$ | 1 |
| Young's modulus of elasticity | $E$ | $\frac{\Delta x^2 \rho_c}{\Delta t^2}$ |
| Constants of Contact force | $K$ | $\frac{\Delta x^2 \rho_c}{\Delta t^2}$ |
| Contact force | $\mathbf{F}^{con}$ | $\frac{\rho_c \Delta x^4}{\Delta t^2}$ |
| Contact torque | $\mathbf{T}^{con}$ | $\frac{\rho_c \cdot \Delta x^5}{\Delta t^2}$ |
| Velocity of agent | $\mathbf{U}_0$ | $\frac{\Delta x}{\Delta t}$ |
| Viscosity of environment of agent | $\mu$ | $\frac{\Delta x^2 \rho_c}{\Delta t}$ |
| Stokes resistance constants for velocity of agent | $\mathbf{K}^t$ | $\Delta x$ |
| Angular velocity of agent | $\omega$ | $\frac{1}{\Delta t}$ |
| Stokes resistance constants for angular velocity of agent | $\mathbf{K}^r$ | $\Delta x^3$ |
| Constants of adhesion energy | $\varepsilon$ | $\frac{\Delta x^5 \rho_c}{\Delta t^2}$ |
| Adhesion energy | $W^{Ad}$ | $\frac{\rho_c \Delta x^5}{\Delta t^2}$ |
| Adhesion force | $\mathbf{F}^{Ad}$ | $\frac{\rho_c \Delta x^4}{\Delta t^2}$ |
| Adhesion torque | $\mathbf{T}^{Ad}$ | $\frac{\rho_c \cdot \Delta x^5}{\Delta t^2}$ |

Table 3.9.1 The symbol of physical entities and their dimension. Note that for some entities the dimension is 1, which means those entities are scalar entities.

In Step 2 of Appendix B, values are calculated for following parameters. The values are listed in Chapters 4 and 6.

| Name of parameter | Symbol of parameter |
|---|---|
| Unit length | $\Delta x$ |
| Unit time | $\Delta T$ |
| Constant for contact force | $K$ |

| | |
|---|---|
| Constant for adhesion force | $\varepsilon$ |

Table 3.9.2 Parameters that the value to be calculated for Chapters 4 and 6

## 3.10 Discussion

Many values for the physical entities are set by making specific assumptions within a parameter range. This range is fixed by the definition of the entities but chosen value may change within this range. As such, there are various sets of parameters values for the model that should provide physically possible behaviours. Although the values estimated in this chapter are based on physical definitions and equation derivation, I may still need to change them during simulation, on account of the need for particular parameterisations for specific models.

Another key aspect is the balance point between adhesion force and contact force. This point determines the amount of elasticity in cells in the model. This point is based on force analysis of two identical cells, and for two non-identical cells I may observe different phenomenon, and this is not considered in the model. Also, when more than two identical cells are sufficiently close, adhesion and contact forces will balance at closer distances because the total adhesion force on each single cell is larger than the situation I analysed. These enhancements would require substantial model development time to address.

# Conclusion

In this chapter a physical model to describe cell-cell interaction and cell-substrate

interaction is presented. Dimension analysis of the equation set and estimation of values for constants is also detailed. This model provides a basis for exploration of cell-cell interactions. Since the model is agent-based, I can model the interactions of individual cells, simulating movement of cells for a certain amount of time (this time is decided by which phenomenon I want to simulate). In addition, by modifying the value of parameters, I may control the cells to produce different phenomena, which makes the model a tool to explore the relationship between physical interaction and morphological patterns that may be formed by cells. In the Chapter 4, 6 and 7, this model is used to study two biological systems: early stage of vessel formation and cancer cell growth.

In summary, the model construction assumes the following:

| Type of assumption | Assumption | Reason |
|---|---|---|
| Geometry of cell | Cells are represented as ellipsoids. | R. Sodt, *et al*, 2014. Hayashi *et al*. 2012 |
| | Cell size and shape do not change during the experiment. | |
| | Mass is evenly distributed within a cell. | For geometric simplicity |
| Force | Only contact force, adhesion force and | Grover *et al*. |

| | | |
|---|---|---|
| | resistance force are significant forces on the cell (density of cell is simiar with water thus gravity does not need to be considered). | 2011 |
| | Contact force can be computed using the Perram-Wertheim approach. | John W. Perram *et al*. 1996 |
| | Adhesion force can be modelled as a step function on distance using Hertz formula. | Ramis-Conde *et al*. 2009, Ramis-Conde *et al*. 2008 and Galle *et al*. 2008 |
| | Contact and adhesion forces balance at a defined point when cells are in contact, and the strengths of the forces can be calibrated based on this. | Based on the fact that cells do not totally overlap |
| | Each cell is considered under effect of resistance force which occurs when cell moves within the fluid. | Beysens *et al*. 2000 |
| | Resistance force can be computed using Stokes' law, and the known properties of the fluid medium. | Brenner, H. 1963 |
| Kinematics | Cells move at very low speed, so their acceleration approaches zero and the forces | J. P. Rieu, *et al*. 2000. G. M. |

| | | |
|---|---|---|
| | upon them are balanced. | Walker *et al*. 2005 |
| Cell-substrate interaction | Interactions with the substrate can be modelled as interactions with mirrored cells. | Similarity with identical cell-cell interaction |
| | The interaction between modelled cell and plane is much larger than interaction between pair of two cells. | Scianna *et al*. 2013 |

# Chapter 4 Vessel Formation: Simulation and Result

## Introduction

Along with content of Chapter 5 and 6, the content of this chapter is part of my journal paper published on Natural Computing. In Chapter 3 a generic model is presented to describe the physical interaction among ellipsoidal objects. The next step is to use that model to study specific biological systems in specific contexts. In order to fit the (generic) model to these systems, parameters may need to be adjusted. Additionally, and depending on the characteristic of the system, more components may need to be added, i.e. processes and associated parameters, into the model. Specifically there is no biological processe in the model: of course, different biological systems have different biological processes.

An important feature of the systems that I seek to model is that they can be described by large number of agents that are controlled by same underlying physical rule set and can present measureable physical behaviour at the system scale. In outline, the approach is to first develop computer code to implement and then execute the simulation, and then find a method to compare the simulation result with experimental data.

The first target system of study is vessel formation. In this system large numbers of the same type of cells interact to form an emergent and particular pattern that is sustained over a long period of time. I measure the position of cells during the pattern formation, and compare the shape of network structure with experimental structures. Changing the parameters in the model affects the shape of network and this can help explain how different behaviours of cells form different vessel formations.

After initial simulations revealed that the behaviour of cells described by my model could, given particular parameterisation, lead to a range of emergent structures including biologically plausible network structures, I reached the stage of calibrating the model with *in vitro* experimental data.

In this chapter, I first discuss the vessel formation simulation, including the impact of different parameters on the emergent structures generated by the simulation.

Microvessels are formed as an emergent structure within the body by the aggregation of endothelial cells, which themselves are formed by differentiation from stem cells. This formation process has three stages (Vempati *et al*. 2011):

- Cell migration and early network formation;
- Network remodelling, where cells connect to each other;
- Further differentiation into tubular structures.

Vessel formation experiments such as (Serini *et al*. 2009) typically use endothelial cells and although these cells do change shape and size during the whole process, during the first phrase, which takes place between six and nine hours *in vitro*, the cell shape and size are the same (Scianna *et al*. 2013). In current literature, the experiments are undertaken in petri dishes and over the course of the experiment cells form 2D net-shaped structures.

Thus vessel formation experiments have a number of features that inform calibration of my model:

- There is only one type of cell, and while the model can simulate behaviours of combinations of cell types, systems of single cell types are an ideal first system;
- The cells maintain the same size and shape over the timescales I are interested in;
- The cells attach to and grow on a (largely) 2D substrate.

Additionally, in the first phase of such vessel formation experiments, cells evidence polarity by aligning in particular directions. Vessel formation experiments explore how the physical interactions among cells, and their low-level physical properties, affect the larger-scale structural patterns in the resulting capillary network. The effects of varying concentrations of growth factors - which have a direct effect on the low-level physical properties of the cells - are of particular interest. In summary the

vessel formation experiment is a good candidate system for model calibration.

Modelling this kind of system has value in understanding the underlying biology. Biological interaction among cells is not only hard to observe but also hard to measure. Physical properties such as cell position and orientation are relatively easy to obtain through imaging, and are linked to biological mechanisms including signalling pathways that control cell-cell interactions. Therefore by mapping the simulation of a model to experiments I can analysis the physical interaction among cells and so inform the connection between molecular control mechanisms and tissue-based structures.

## 4.1 The *In Silico* Experiment

### 4.1.1 Simulation Process

The implementation of the vessel formation simulation using the model in Chapter 3 follows the structure shown in Figure 4.1.3.1. I have chosen an agent-based modelling approach. This allows me to define interacting rules for single cells, and examine both the lower-level properties of individual cells and the higher-level behaviour of the system as a whole.

In the simulation, all the agents are stored in a list. Firstly an agent is chosen from the list; then the simulation calculates the contact force between the chosen agent and

each other agent in the list. After that all the contact forces are added up to produce the total inter-cell contact force for each cell. The total inter-cell contact torque, total inter-cell adhesion force and total inter-cell adhesion torque are calculated in the same way.

Then the simulation calculates the physical interaction between the chosen agent and the substrate. All the forces are added up to calculate velocity according to Stokes equation, and all the torques are added up the calculate angular velocity. At last, according to the velocity and the original position of the chosen agent, its position after 1 unit time is calculated; according to the angular velocity and the original orientation of the chosen agent, its orientation after 1 unit time is calculated.

Each agent in the list is chosen one by one to calculate and update its position and direction. The process is repeated until the desired process time is reached. The flow chart of simulation is shown in Figure 4.1.3.1.

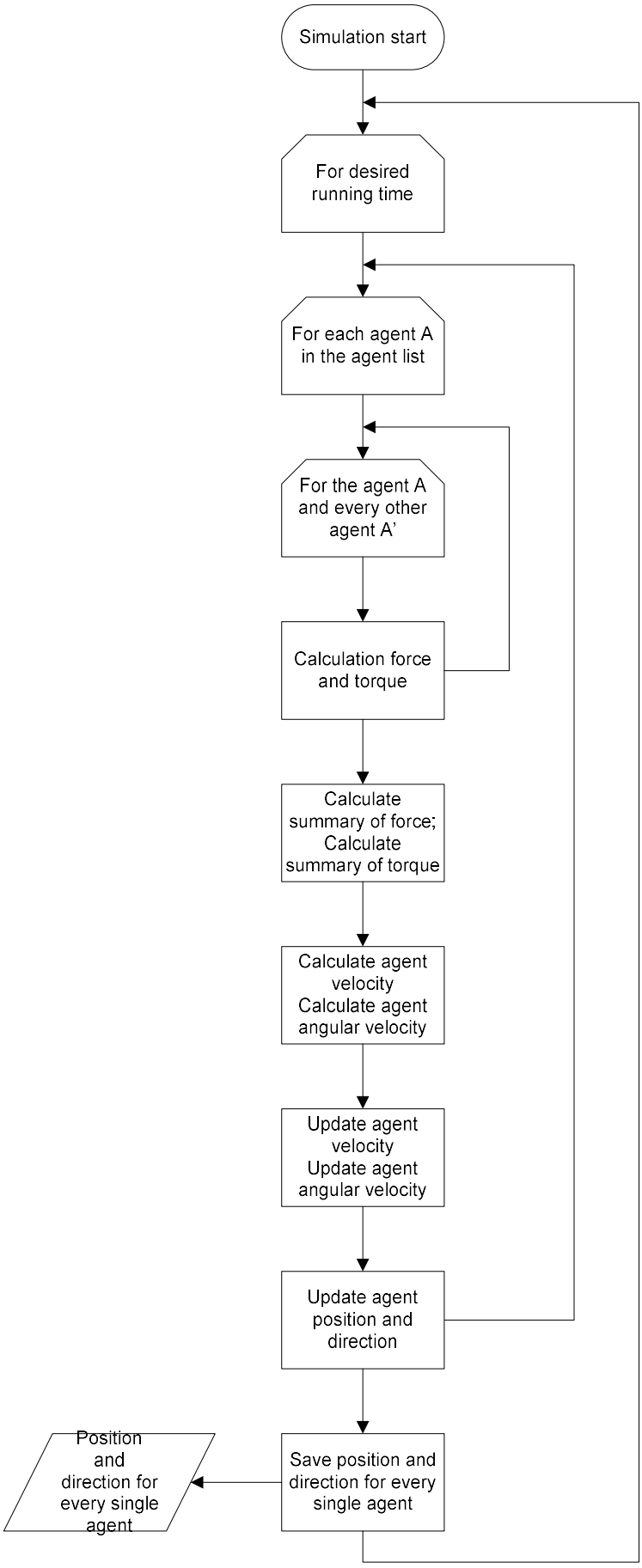

Simulation start
For desired running time
For each agent A in the agent list
For the agent A and every other agent A'
Calculation force and torque
Calculate summary of force; Calculate summary of torque
Calculate agent velocity Calculate agent angular velocity
Update agent velocity Update agent angular velocity
Update agent position and direction
Save position and direction for every single agent
Position and direction for every single agent

Figure 4.1.3.1 Simulation flow chart

The agent-based model requires a large number of objects to show distinct emergent behaviour. At the start of the simulation, 2000 agents are randomly placed on a horizontal plane. This number is estimated from the 'middle' number density of cells in (Serini *et al*. 2003). The agents are represented by identical ellipsoids with length of three semi-radius satisfying 1:1:4, as shown in Figure 4.1.3.3. This ratio is chosen because this long and thin shape helps to show the polarity property of cell; at the same time, it is not too long to be unrealistic comparing to real cells observed in experiment, as shown in Figure 4.1.3.2.

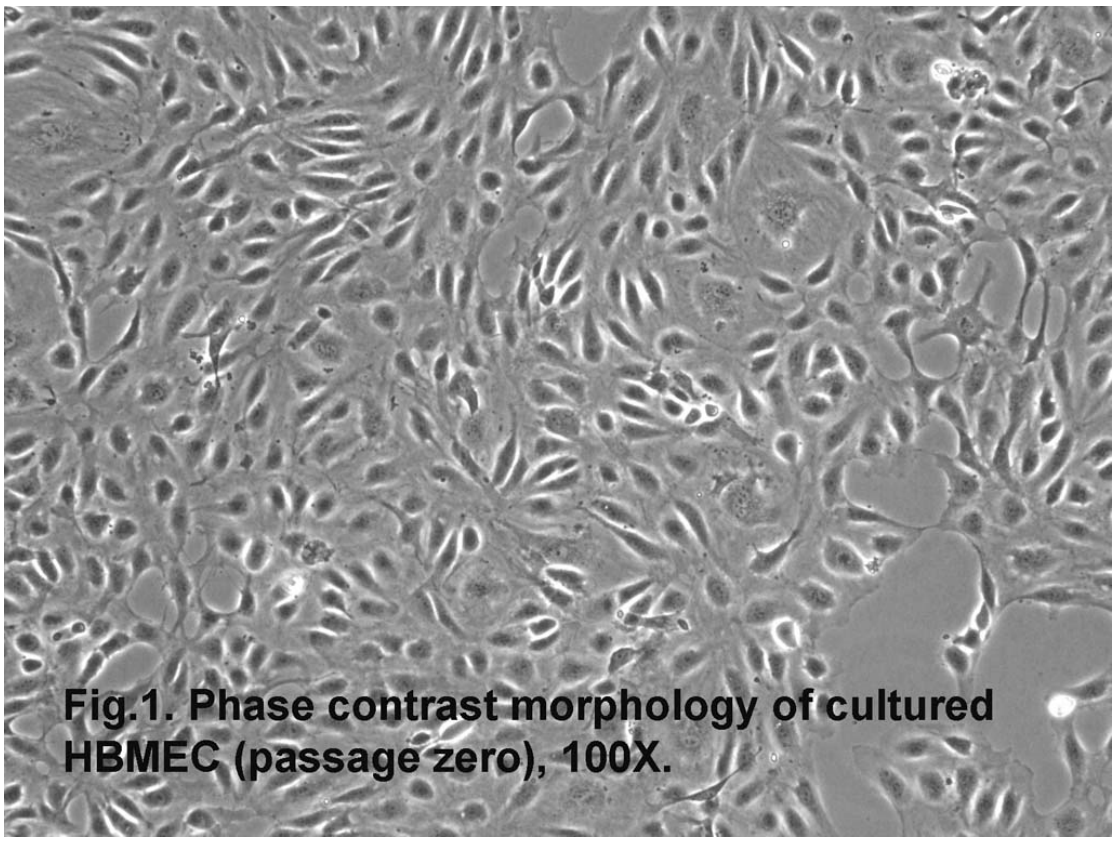
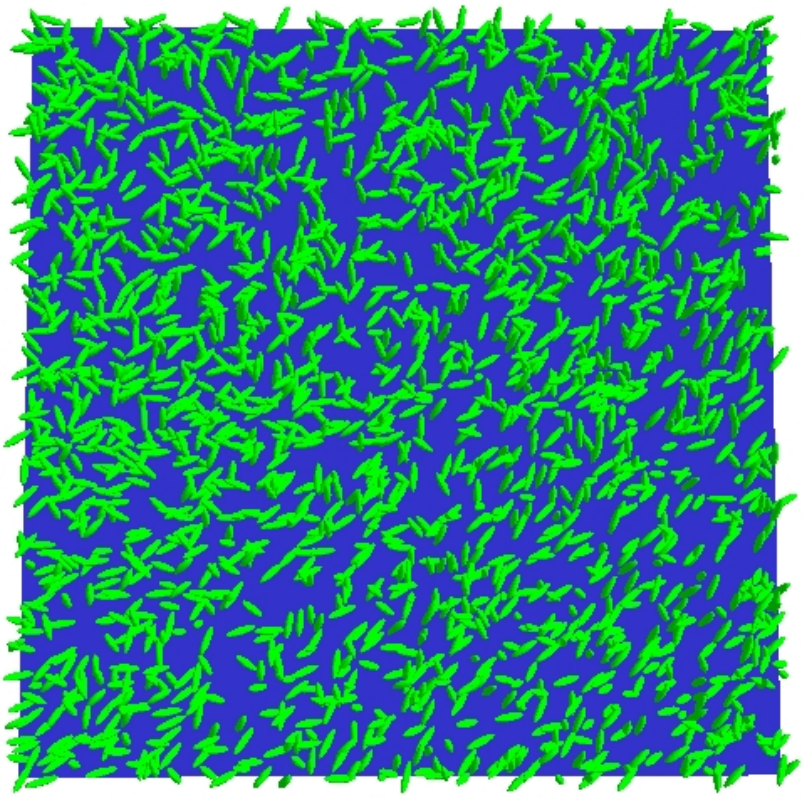



Figure 4.1.3.2 Shape of Human Brain Microvascular Endothelial Cells [http://www.sciencellonline.com/site/productimages/large/1000-1-l.jpg].



4.1.3.3 Starting point of a large-scale simulation, in which 2000 agents (green) are randomly placed on substrate plane (blue).

The only restriction on the starting position of cells is that they cannot be too much closer than the distance where the contact force and adhesion force balance in order to allow a realistic balance between contact force and adhesion force. This distribution is implemented in a simple manner: the horizontal position of cells is generated for one cell at a time, its value is randomly distributed within the substrate; the vertical position, on the other hand, is assigned a value that is slightly longer than the shortest semi-axis of agents. The setting of the vertical (with respect to the substrate) position of agents is to simulate the attaching behaviour of *in vitro* cells to the substrate.

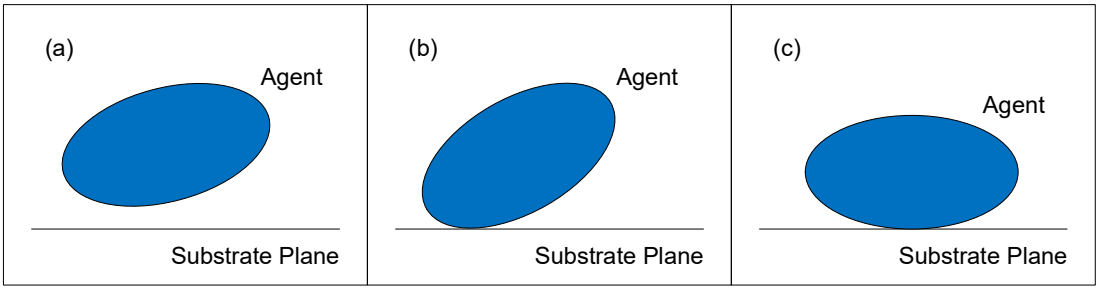



Figure 4.1.3.4 (a-c) The interaction between an agent and the substrate. Firstly the agent rotates (anti-clockwise) and moves towards the substrate until they contact each other (a-b), then the agent rotates (clockwise) and moves until the longest semi-axix is parallel with the substrate plane (b-c).

The settings of positions are to ensure that the longest semi-axis of agents is horizontal to the substrate and all the agents attach to the substrate along the shortest semi-axis shortly after simulation starts (as Figure 4.1.3.4 shows). The direction of cells is also random. I save to a file the above mentioned geometry information for each cell separately and store them in a list, which I denote as the 'cell list' in the following discussion, i.e. there is a list in program and each record in this list represents the geometry information of one cell. I have 2000 agents so that the list starts with 2000 records.

## 4.1.2 Update physical information of cell

When the simulation starts, the program loops through all the cells and for every single cell calculates the contact and adhesion force and torque between it and all the other cells. The program keeps record of the sum of all the contact forces, contact torques, adhesion forces and adhesion torques for each cell, and when the looping is completed for all the cells, the program loops through all the cells again, this time to

calculate its interaction with the substrate.

As discussed previously, the interaction between a single agent and the substrate is considered as the interaction between this agent and its mirror image on the surface of the substrate. In the program I do not in reality generate a mirror image and then add it to the cell list; instead, I calculate the geometry information of the mirror image and pass it into functions that calculate interactions, and then only store the calculation result. The forces between one agent and its mirror image are added together with the sum of all the forces, and the torque added to the sum of all the torques. The summary force and torque are used to estimate the velocity and angular velocity of the agent, and then with its original position and velocity the new position of agent is calculated.

Similarly the original orientation and angular velocity are used to generate the new orientation of the agent. Note that there is an implicit entity in the calculation, which is the time step that one position/ direction update takes. The choice of time step size (i.e. unit time in the model) is an important factor. The time step must be short enough to obtain results at a comparable temporal resolution to the *in vitro* experimental data. However, smaller time steps require more calculation steps to simulate the same length of real-world time. In 3.10 I estimated the value of this time step to be one second. The one-second time step results in *in silico* experiments that take an impractically long time to run. It is necessary for model and simulation testing because I can observe the movement of agents with more detail, even though each test

run takes up to 120 hours to generate result over biologically meaningful timescales.

The physical information of agents is saved into text file regularly; each time the program generates one new file for all the cells to save their current physical property. As I assume the size and shape of cells do not change, I simply need to keep record of cell position and orientation. For testing purposes I also keep velocity and angular velocity for each cell. The interval between data recordings can be adjusted by changing the code of simulation program. Obviously, the shorter the interval is, the larger the files to be generated, since it provides more detailed description of cell interactions.

Data files can be imported post-simulation into a separate visualisation program when the simulation is completed. In this way I have an intuitive representation of the shape of structure formed by cells. As the coordinate system in the visualisation framework is different from the one I are using in model, a coordinate transform is required when cell position data is imported. This part was tested beforehand to make sure the cells are at correct place with correct direction. The visualisation can be saved as imaged files for later use.

### 4.1.3 Simulation parameters

Simulation parameters are also calibrated as described in 3.10. However, early test simulations showed that the agents moved unrealistically rapidly, and this

demonstrates that agents are affected by overly large forces. It does not mean the calculation of forces or torques is incorrect, but to investigate their motility the agents needs to be slowed down. Thus the inter-agent interaction needs to be reduced. Because the contact force only takes effect when agents contact each other, the adhesion force is the main cause of agent motion. Therefore the first step is to reduce the size of the adhesion force. However if the contact force remains the same while adhesion force is decreased, the position where the two forces balance is changed, which means the elasticity of modelled cell is changed. To maintain the elasticity of modelled cell, the contact force and adhesion force should be reduced by the same scale.

By running simulations with various values of forces, it is observed that agents bounce away with high speed when they make contact. A plausible explanation is that the high movement speed of the agents causes a large repulsive contact force and then a high speed in the opposite direction. Thus I determined that if both forces are reduced by an order of magnitude, the movement speed of agents is reduced to a plausible value. Considering the simplifications used during dimensional analysis (see Section 3.9 and Appendix B for full details), the oversized forces might be caused by the errors in estimation process of constants in equations. In the estimation process, for several times the shape of cells is assumed to be a sphere, however in the simulation, the long and thin ellipsoidal shape is used. In other words, the methods used to calculate forces are not wrong. To solve the problem, the constants in contact

and adhesion force equations are reduced by an order of magnitude. Clearly, with access to experimental data sets a fuller calibration of these values would be possible.

| Name of parameter | Value |
|---|---|
| Unit length $\Delta x$ | $1.25\times10^{-6}$ m |
| Unit time $\Delta T$ | 1s |
| Density of cell $\rho_c$ | $2000kg/m^3$ |
| The proportion of typical cell size and unit length | Three semi-radii of ellipsoidal agent: $5.0\times10^{-6}$ m, $1.25\times10^{-6}$ m, $1.25\times10^{-6}$ m |
| Dynamic viscosity of fluid in which cells are moving $\mu$ | $1\times10^{-3}Pa\cdot s$ |
| Constant for contact force $K$ | $6.4\times10^{10}$ |
| Constant for adhesion force $\varepsilon$ | $3.2768\times10^{10}$ |

Table 4.1.1 Value of parameter used in vascular formation. Note that the values of $K$ and $\varepsilon$ are their dimensional value $\tilde{K}$ and $\tilde{\varepsilon}$. As $\tilde{K}$ and $\tilde{\varepsilon}$ do not appear in Chapter 3, I use $K$ and $\varepsilon$ for a better understanding

## 4.2 *In Silico* Experiment Result

### 4.2.1 Agent movement

It is observed that agents tend to stick to each other at the longest end (as shown in Figure 4.1.3.6), as in biological systems. Note that there is not biological setting in the model to drive this behaviour of agents. In my model, in addition to adhesion force there is adhesion torque, which drives nearby agents to change their directions while moving together. Additionally, the equation to calculate adhesion torque contains a term representing contact potential, which is multiplied by another term representing the distance of agents. For two fixed agents, the distance is also fixed, but the contact

potential is smallest when the two agents have the longest semi-axis pointing to each other. In other words, the adhesion torque reaches its smallest value when two agents are pointing to each other along the longest semi-axis. Any other angle produces a higher adhesion torque, which turns both agents to the direction that produces lower adhesion torque. This phenomenon is an emergent behaviour of pure physical interactions which are determined by shape and positions of pair of agents. I can see that physical rules can drive model to shows biology phenomenon such as the anisotropy of modelled cells.

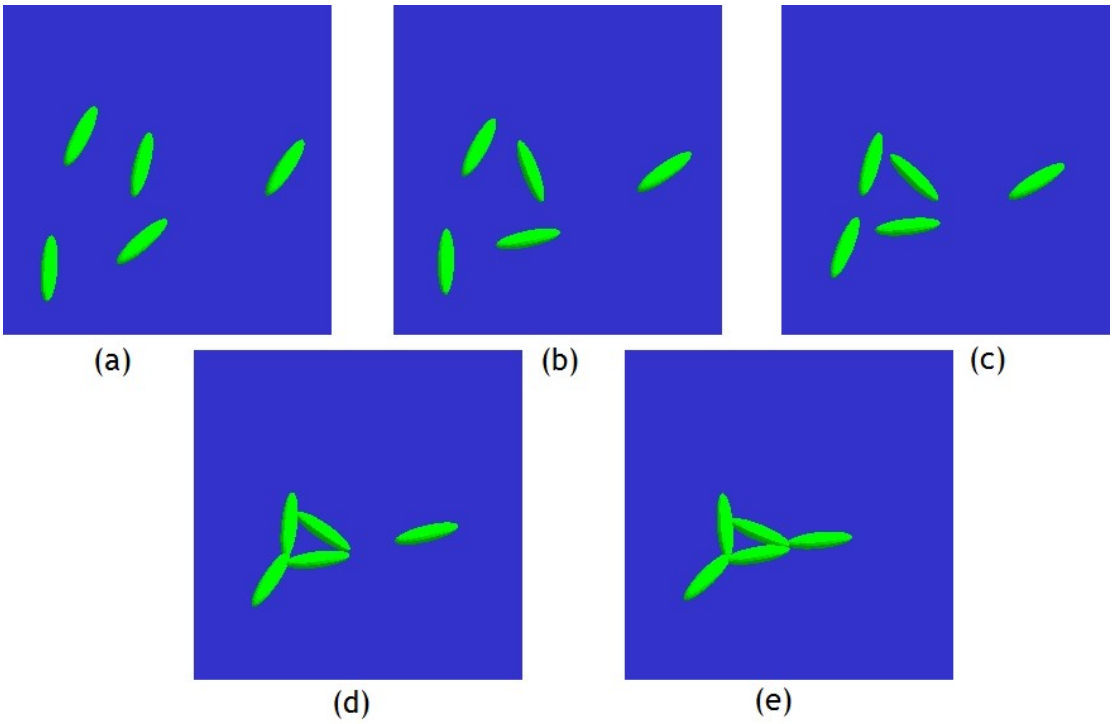



Figure 4.1.3.6 (a-e) In the simulation, agents turns to each other and stick along longest semi-axis

## 4.2.2 typical local patterns

According to this feature outlined in the previous paragraph, there are three typical types of patterns, as shown in Figure 4.1.3.7. Type 1: in the areas where agent density

is high, agents may be able to connect with several other agents and form a star-like structure. Type 2: at intermediate densities, agents are close enough to contact each other but not close enough to form the type 1 pattern; instead, agents only connect with their nearest neighbours at both ends and form a line. Type three: in regions of low density, the distance between agents are too large so that agents cannot connect with each other but are attracted to nearby higher density area or remain as separated agents. Together these three types form a net-shaped pattern: the type 1 as the junction, the type 2 as the branch and the type 3 as the holes of the net.

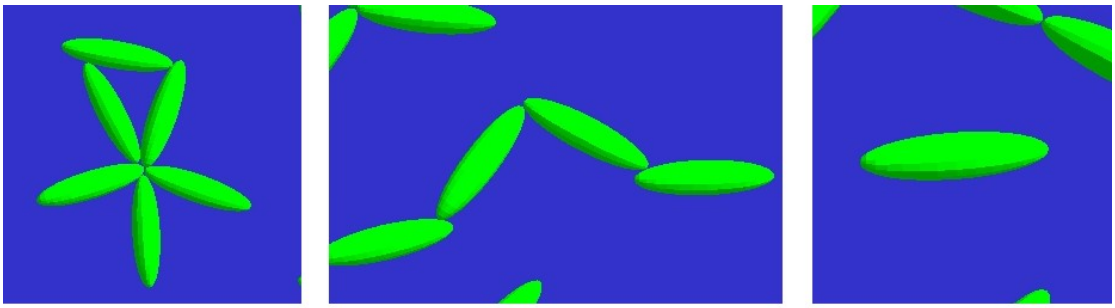



Figure 4.1.3.7 Three typical connect type of agents: (a): ‘star’ type; (b) single connection type; and (c) non-connection type.

## 4.2.3 overall pattern

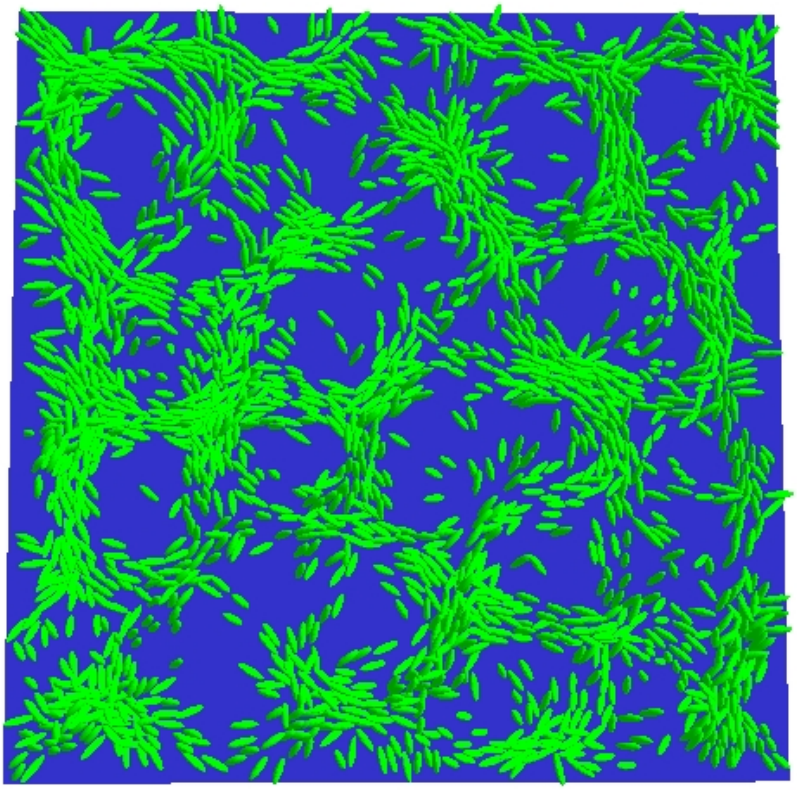


Figure 4.1.3.5 Network structure formed by agents (in green) on substrate plane (in blue).

Over the course of the simulation, agents gradually move towards other nearby agents under the effect of adhesion force resulting in the pattern shown in Figure 4.1.3.5. The balance of adhesion force and contact force eventually stops movement. At the balancing point agents are partly overlapped, and the degree of overlapping shows the level of elasticity of agents.

4.3 result analysis

The pattern generated by simulation may be compared with the result vascular network of *in vitro* experiment shown in Figure 4.1.3.8: in both structures, there are junctions composed by group of cells (agents), and the junctions are connected by

lines of cells (agents); between the connected junctions there are holes where separate cells (agents) exist. Both structures have net-shaped feature. Similar with cells, the agents tend to connect with each other along the longest semi-axis. The analysis of the measurement of structures is in section 4.1.4.

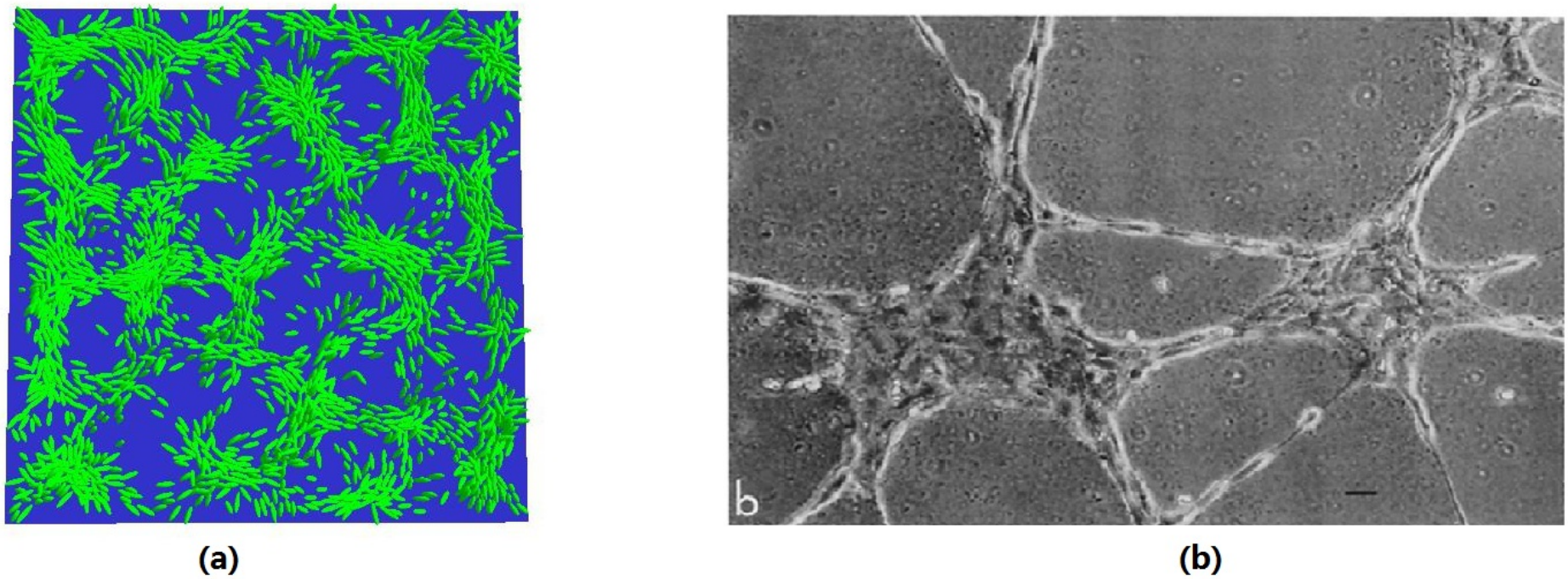



Figure 4.1.3.8 (a) structure generated by simulation; (b) structure generated in *in vitro* experiment (Kubota *et al*. 1988)

## 4.2.4 pattern stability

However at this stage, the model is not calibrated with any experimental data, resulting in the adhesion force being much larger and having an effect at longer distances than it should. Agents interact in the simulation, for a period equal to 15 minutes in the real world, to the form structure shown in Figure 4.1.3.8 (middle). Further, as the simulation progresses the nearby cells will gather up and form cell mass and the net structure is destroyed. If I allow the simulation to continue past the state shown above, the pattern will collapse into a few large clusters of cells, as shown in Figure 4.1.3.8 (right). Thus the observed network structure is transient and not stable.

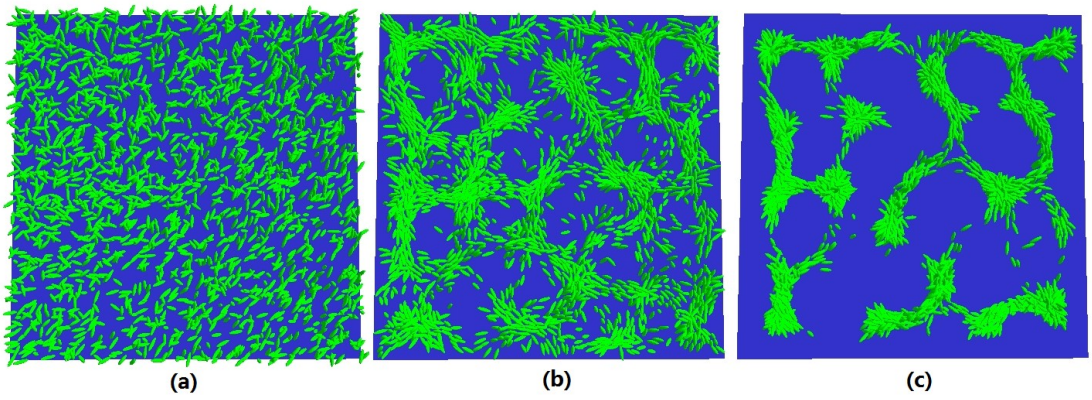



Figure 4.1.3.8 (a) - (c) The starting, middle and end point in of simulation. The structure is formed equal to 15 minutes in the real world from (a) to (b). The structure is collapsed in (c)

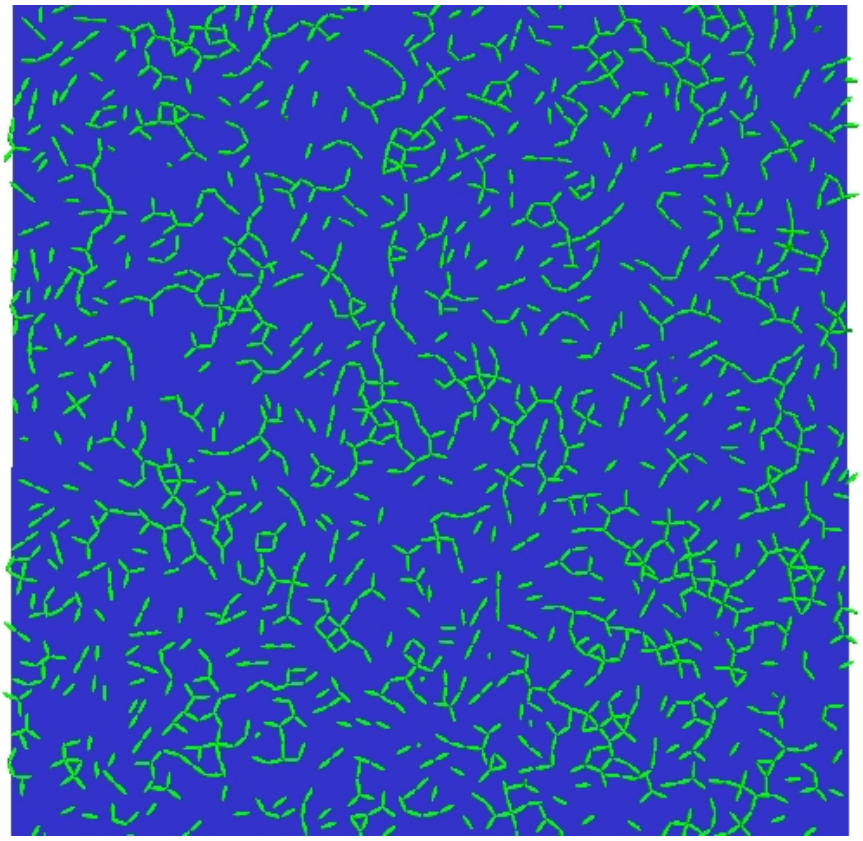


Figure 4.1.3.9 structure formed by adhesion force that has a finite range of effect

To generate a stable structure, I have to assume that the adhesion force has a finite range of effect. In this way an agent only connect with a limited number of neighbours. The adhesion force amongst one single agent and its neighbours in one direction may then be balanced by the force amongst it and its neighbours in other directions. Thus the net-shaped structure is maintained.

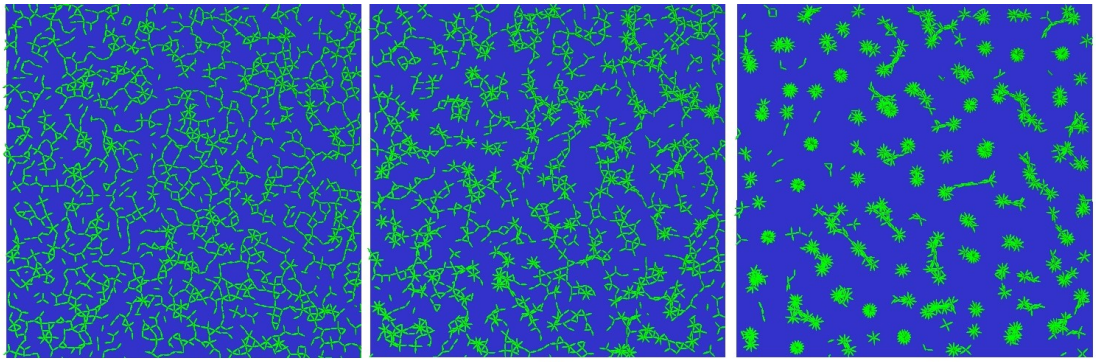



Figure 4.1.3.10 Simulation result of structure formed by 2400 agents with different adhesion force effective range setting. (a) Structure generated with the effective range of ahesion force equals 3; (b) structure generated with the effective range of ahesion force equals 4; (c) structure generated with the effective range of ahesion force equals 8.

By exploring various adhesion force effective range settings, I determined that the best result appeared when contact potential equaled 3. For longer effective ranges, such as 4, the vessel branches are thicker but there are fewer connections between branches; for even longer ranges such as 8, the structure collapses, as shown in Figure 4.1.3.10. With the adjusted model constants, a cell will not adhere to a far away cell, and the contact force and the adhesion force balances at an appropriate distance so that two nearby cells will not overlap too much.

This image also shows that there is no typical net structure in this simulation. It is because the cell density is too low. With more cells, the branch-shaped structure can form holes and net-shaped structure, as shown in the figure 4.1.3.11.

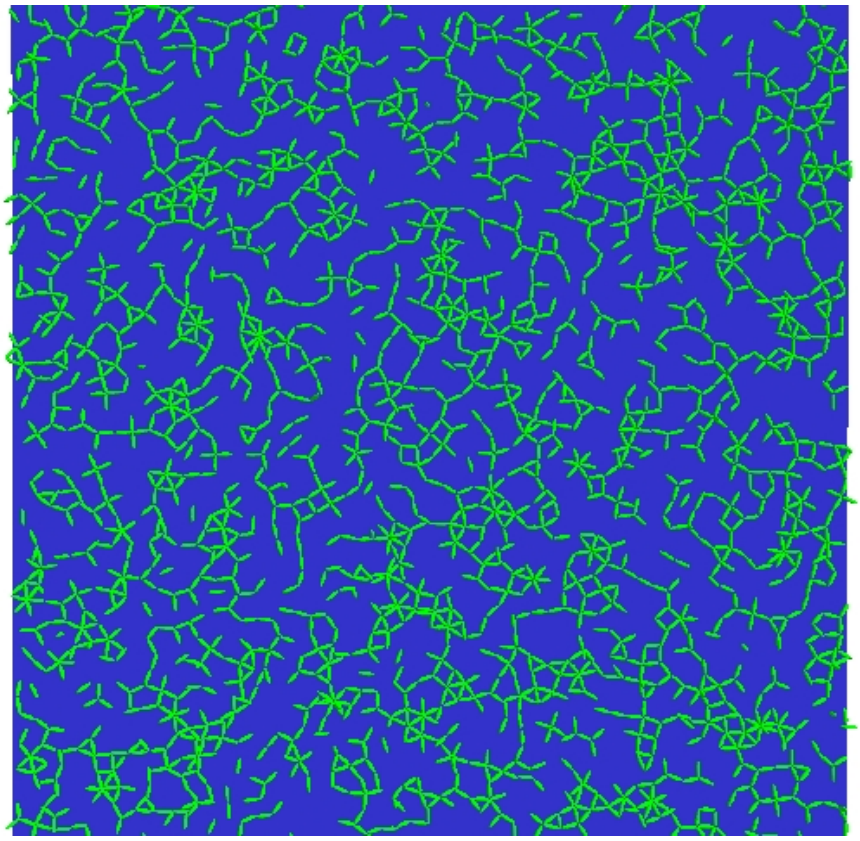


Figure 4.1.3.11 Simulation result with high number density of agents.

This is the simulation result started with 2400 agents. There are more holes formed with single lines of agents. Also in both simulations shown in Figure 4.1.3.9 and Figure 4.1.3.10 (a) and (b) the structure is stable. As the contact force and the adhesion force are balanced, agents are only affected by the random force used to simulate Brownian motion. In the stable stage, the agents slightly shake at their own position and no agent connection breaks.

## 4.3 Result analysis

Two typical simulation results are shown in Section 4.1.3. I can see that the cells quickly moved towards the substrate and attached to the surface, as Figure 4.1.4.1

shows. Then the cells move towards each other and tend to connect along the longest semi-axis (as shown in Figure 4.1.3.6). After the network is formed, the network is stable and keeps the observed formation.

(a) (b) (c) (d)

Figure 4.1.4.1 The blue part is the substrate, from (a) to (d), the green agent moves toward substrate and attach to surface.

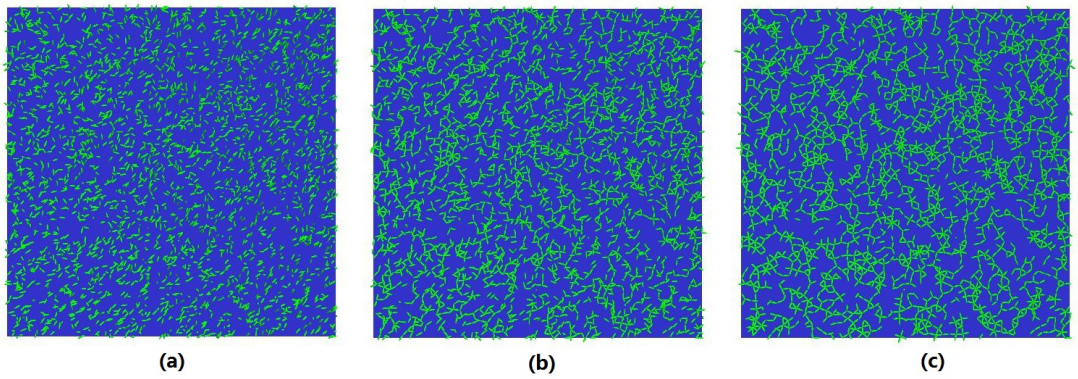



Figure 4.1.4.2 (a) - (c) Screenshot of start, middle and end of simulation

By comparing the *in vitro* experiment result and *in silico* image (as shown in Figure 4.1.3.8), I can see that both results contain a net-shaped structure. The structure can be considered as holes that are surrounded by several cells, and conjunctions that are where cells gather up. Arguably there is a degree of similarity, but I need a quantitative method to measure both structures for further study.

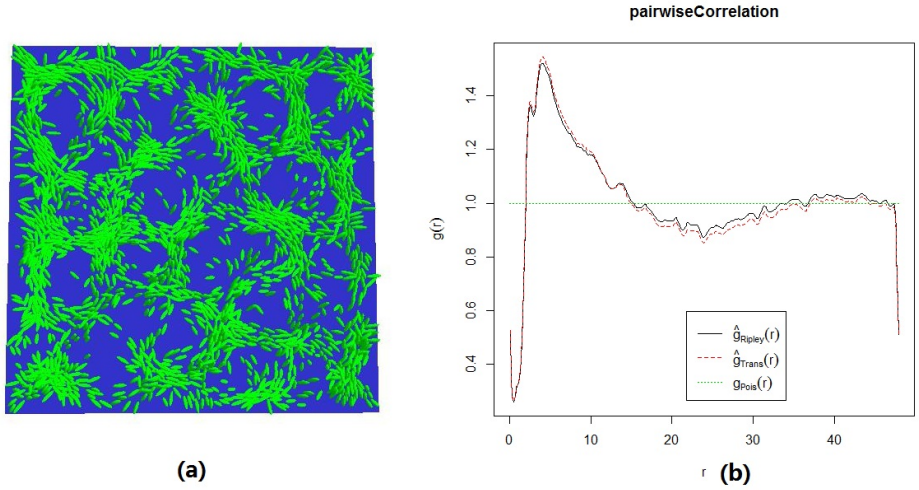



Figure 4.1.4.3 (a) Simulation result of vessel formation; (b) pairwise correlation function of (a)

By using the 'pcf' function 'pairwiseCorrelation' in R [http://127.0.0.1:25934/library/spatstat/html/pcf.html], the pairwise correlation curve is generated, as above diagram shows. In the correlation curves, x-axis represents the

distance between arbitrary pair of objects, and y-axis means the chance that two agents occur at that distance.

Comparing with Figure 4.1.4.3, I can see that both curves have a peak near the origin, which is a value higher than 1, followed by a valley with a value lower than 1; in the rest part both curves fluctuated around value 1. The experimental curve starts with a high value near the origin, while the curve produced by simulation starts with a low value. Note that Figure 4.1.4.3 is generated by a continuum model and Figure 4.1.4.4 is generated by an agent-based model, the early part of the curve will be affected by the coarse-size of cells, which is very different. As discussed previously in this section, a y-axis value above 1 means the objects tend to cluster at this distance; a y-axis value below 1 means the objects tend to scatter, or repel, at this distance. Therefore the x-axis value of the valley represents the typical size of hole in the net-shaped structure. The minimum value of simulation is near r=20. By changing the value of parameters in the model, the position of minimum value and other feature of pairwise correlation curve can be changed. The detail is discussed in section 6.2.1 in the study of cancer cell interactions. Given these broad similarities, and suitable effort to explore the parameter space, including cell size, the simulation results could be better fitted to the experimental data.

In (Serini *et al*. 2003), the increased concentration of growth factor increases the size of hole in the pattern. Unlike this model, there is no growth factor in my model. In my

model, I assume that the adhesion force plays similar role in the pattern formation. I would need to carry out simulation with various parameter settings to verify this assumption. If confirmed, I would also need to verify the quantity relationship between the concentration of growth factor and adhesion force. However as reported in Chapter 1, it was not possible to carry out the necessary and model-led set of experiments, and so I converted the model to a new system where I had data streams available to me.

## Conclusion

In this chapter, simulation process and parameter setting are discussed. The initial results of simulation, of both local and overall pattern, with different parameter settings, are explored. These show that the model has the ability to form higher-scale structure emerging from the interactions among a large number of identical objects controlled by simple rules.

It is also noticed that with certain parameter the net-shaped structure collapse over time, while some parameter settings can keep the shape of structure, which shows the sensitivity of the group behaviour to changes in certain parameters in the model.

The next step is to calibrate the model with experimental data. However, because of the project scope change described in Chapter 1, I did not have the expected access to the source of experimental data. Hence I hope to reuse the model in another context

where I were able to design experiments and gather the necessary data.

# Chapter 5 Predicting population growth curves by mechanisms of drug action and hypoxia

## Introduction

According to Chapter 4, the model is able to describe a system that is made of simple objects and shows different behaviour with different parameter settings. However due to lack of experimental data, I cannot calibrate the model, and therefore want to reuse the model for another system, which has emergent behaviour, and of which I am able to design experiments and gather the necessary data.

In this chapter, model reuse is discussed in Section 5.1. An important step is to determine systematically to what extent the model for vessel formation may be reused in cancer cell simulations. I use the CoSMoS project (Paul S. Andrews *et al*. 2010) to consider in a structured manner how the physical interactions among cells, including the implemented software to simulate it, may be adapted to a different context. As the modes of physical interaction among cells are broadly similar between the two model systems, this seems intuitively to be an appropriate approach - but identifying and revalidating my assumptions using the CoSMoS process both exposes thinking and gives confidence to the simulation's repurposing, and might enable the future reuse of the physical model in other contexts.

From repurposing analysis in Section 5.1, a conclusion is reached that the model can be used to describe cancer cell growth with cell cycle and cell population growth being modelled and added to the initial model. Then the cell cycle model to be added in my model is discussed in Section 5.1.1, and the population growth model is introduced in Section 5.1.2.

Then with was collaboration between the author and an experimental biologist, a group of experiments are designed to calibrate the cell cycle model and population growth model (Section 5.2). I am interested in the effect of 5-FU, hypoxia and combination of 5-FU and hypoxia on cell growth. In order to model the 5-FU and hypoxia mechanisms, the first experiment is to measure the control population as the base line. In line with this, the model is first calibrated with control population. Then by considering the biological function of 5-FU and hypoxia, a plausible mechanism is added to the model for each intervention. The model initially calibrated for the control population is then able to predict the dynamics of the population in the presence of 5-FU by only invoking that mechanism associated with 5FU. Similarly the mechanism of hypoxia is added and the population curves successfully predicted. At last, for independent validation of the mechanisms added for both 5-FU and hypoxia effects, both mechanisms are set to work together and simulation result is compared with the experiment of population growth under the effects of combination of 5-FU and hypoxia. Note the model is not calibrated against that combination experiment.

These experiments are carried out by out cancer cell biology experts. The experimental result is shown in section 5.3, and the detail of population growth calibration is discussed in section 5.4.

The experiments are carried out on petri dish, which means the cells can be considered as growing on a 2-dimensional plane. The simulations are done with agents on a 2-dimension plane as well (the agents are ellipsoids moving on 2D plane). In future I plan to expand this model to simulate cell growth in 3D space, and the model implementation allows this. The study in 2D space can be considered as a special condition of 3D growth, which is normally referred to as spheroidal growth. It is reasonable to assume that the growth processes in 2D and 3D have some commonalities. In fact (Radszuweit, M., et al. 2009) suggests that for agent-based models, 2D growth simulation is sufficient to describe generic features of population growth, unless the shape-specific data is required (from simulation). Therefore the study of 2D population growth is important and its cell-cycle parameters may be used in future spheroid growth simulations.

## 5.1 Model Repurposing

The existing physical model has already demonstrated the ability to reproduce spatial patterns of cell growth resulting from physical interactions within an agent-based simulation, and I have existing tools to visualise and analyse the output from the model. I would like to reuse as much of this infrastructure as possible to build on

previously tested work and so maintain simulation quality and reduce development time. I introduced concepts from CoSMoS modelling process to analyse and compare two models, please see (Ye Li *et al*. 2014) for detail. Here I just repeat the summary. I first need to identify the parts of model that I can keep unchanged. Then, through analysis of requirements for cancer cell growth study, I list the parts of the model that need amended, including reconsidering (and re-evaluating if necessary) previous assumptions, as well as new features that need to be added into model.

Although the behaviour of cancer cells is different from that of endothelial cells, each is system of cells of the same type that are controlled by the same set of physical interaction rules. For this reason I keep using agent-based modelling process. I consider cancer cells as ellipsoids as well, so that the method to calculate physical properties can be reused, except some of the parameters needs re-evaluated. The value of some parameters can be obtained from *in vitro* experimental result, as discussed in Section 5.2.

There is one parameter, however, that does need re-evaluation but cannot be obtained from analysis of experiment results, which is the time step in the simulation. As the growth and movement of cells are very slow, and more importantly, I are not able to measure cells very frequently (see next section) the time step, i.e. $\Delta T$ in Chapter 3, has requirement to be less than 10 minutes, while in simulations in Chapter 4 the value of this parameter is set as 1 second. Since this is one of basic unit entities I need

to estimate the values of other constants, in the new model those other constants will need to be adjusted in line with this change in time step. According to ‘Step 1’ in Appendix B, the terminate velocity is within a constant range regardless of the value of $\Delta T$. Thus $\Delta T$ can be increased from 1 second to 10 minutes, with increase of the unit length $\Delta x$.

Because cancer cells grow and die/divide, my assumption that cell size and shape do not change, is no longer valid. This requires to add a couple of new mechanisms to the model. One of these mechanisms is the cell cycle, in which agents gradually change their shape or size, and all the forces and torques are calculated according to the changing shape and size. When the cell reaches a certain age, it may die or divide into two daughter cells. The chance it that dies, termed death rate, is the second mechanism to be added. Since I are using the agent-based modelling process, the death rate is added as a rule affecting all the agents. By affecting the chance that every single agent duplicates itself, the death rate affects the population of a group of agents over time. These important additions are described in Section 5.1.1 and 5.1.2.

## 5.1.1 Modelling the cell cycle

The cell cycle model part is independent from the physical part of the model, and thus can be calibrated separately and is the focus of this Chapter. It is also necessary to calibrate the cell cycle model first based on population growth data, as to test the

physical part of model, the spatial analysis is required to analysis the pattern generated from model simulation, for which the population dynamics of the agents is crucial (see Chapter 6).

Cells can be considered as state machines (Long *et al*. 2012). As discussed in 4.2.2, to simulate cancer cell growth and division I need to include a cell cycle into the model, because during this time the cell will change its size, and this may affect the physical interaction. A typical cell cycle includes 3 phases: quiescent, interphase and cell division (Geoffrey M Cooper, 2013). Summarised for convenience here, and see section 2.4.1 for more detail on the cell cycle, in each phase the cell complete certain tasks and has to pass a checkpoint to go to the next phase. The quiescent phase, or called $G_0$ is a resting phase. The interphase consists of three sub-phases: $G_1$, $S$ and $G_2$. The cell increases its size in $G_1$, then it has to pass $G_1$ checkpoint to get into phase $S$, in which the DNA is replicated. Then the cell gets into $G_2$ and continues to increase its volume, and passes the $G_2$ checkpoint to get into the mitosis phase, also called the cell division phase.

Importantly the biological activity of the cell is not the focus of the model here. I only consider the external phenomenon of cell behaviour, regardless of its internal mechanisms. I simply need to account for different physical phenomena, such as increase in cell volume, regardless of the underlying biological mechanism, such as replication of DNA. So this biological cell cycle is abstracted to cell stages in my

model, which are corresponding to external, physical cell behaviours.

Thus in the model there are four stages: a Stable stage, in which cell does not change at all; a Grow stage, in which cell gradually increases its volume; a Split stage, for cell proliferation, in which a cell divides into two daughter cells; a Die stage, in which cell dies and disappears. These are simplifications of cell behaviour as described in section 2.4.1 and may be linked to underlying biological mechanisms in the way that they change the phenomenological model. For example, in real systems cells will not simply disappear when they die, some of the cells gradually shrink; while others expand and at last explode (Kerr *et al*. 1972). By simplifying the cell cycle to four phenomenological stages, I can focus on the physical properties of and interaction among cells, which is the key focus of the modelling.

The following state machine diagram represents the four stages and the transitions between them. The stages are in form of circles (simple circle, double circle or black edged circle) and the transformations are represented as arrows from start stage pointing at end stage.

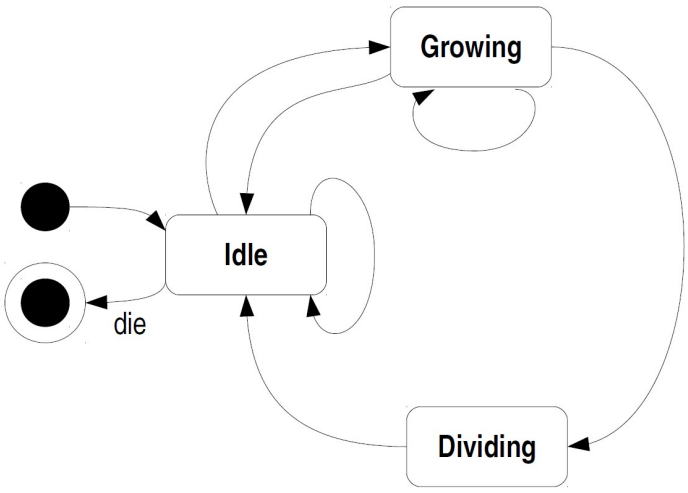



Figure 5.1.1 Simplified cell cycle of agent as a state machine (Ye Li *et al*. 2014). The cell cycle starts from black dot and ends at white dot with black centre. Blocks are states which represent the stage of cell cycle, and arrows show the possible turnover of stage. For example, a cell can turn from Growing stage to Dividing stage because there is an arrow pointing from Growing block to Dividing block. But there is no arrow pointing from the Dividing block to the Growing block, which means cells cannot turn from Dividing stage back to Growing stage.

Figure 5.1.1 shows the state machine that models a simplified cell cycle and drives the behaviour of the simulated cell. This represents the observed behavioural modes of the cell – quiescence, growth, reproduction and apoptosis. The black circle is the start state. It means that every agent starts then turns to the Idle stage at once. Then the arrow pointing from the Idle status to itself means that an agent may decide to stay in the Idle stage, corresponding to the cell quiescence in the real system. The agent may also decide to switch to the Growing state, corresponding to the cell growth in the *in vitro* experiments. This switch is represented by an arrow from the Idle stage to the Growing stage.

As an example of reflecting biological processes in a phenomenological model consider a particular intra-cellular signalling pathway called Hippo. Environmental

signals such as local cell density may be regulated by signalling pathways such as Hippo. The cell can detect its local environment by membrane receptors: if the cell detects limited space in the surrounding tissue, the Hippo pathway is triggered to block the transcription of gene in the cell nucleus (Liu *et al*. 2012). This process is also called contact inhibition, and is why in the model I keep a switch from the grow state to stable state. Cancer cells, on the other hand, may have dysregulation in this pathway and so are not controlled by the contact inhibition, so by turning on/off this switch in the model I can simulate cells with normal/cancer contact inhibition.

When an agent is in the Growing stage, the agent may remain in this state, which is represented by an arrow starting and ending on the Growing stage. This relates to agents increasing in volume in the *in vitro* experiments. In the model, when an agent stays in the Growing stage; it gradually increases its volume. I assume that the volume of an agent increases linearly in the Growing state. The rate of increase is calculated by measuring the change of volume and total time spent in the state. As discussed above, the agent may switch between the Idle and Growing state, and this defines the total growth time, in which the agent doubles its volume.

When an agent grows to twice its volume, it will transfer from the Growing stage to the Dividing stage. This is a stage that agents cannot remain in: every agent that is transferred into this stage will be divided into two daughter agents and both daughter agents will transfer to the Idle stage at once.

Finally, there is a typical length of time an agent can live, which is referred to as ‘typical agent age’ in the following sections. As agents may stay in the Idle stage for a period of time, or switch between Idle stage and Growing stages several times, agents may reach this typical age without being fully grown. In this case the agent will die, which is indicated by the arrow starting from the Idle stage and ending at the die stage (white circle with black dot in the centre). I also need to accommodate the fact that in real experiments the time each individual agent can live may vary, so that I assume that the real cell age is normally distributed with the typical age as the expected (mean) value with an appropriate variance (Standard deviation).

In general an agent cannot outlive its typical age, except if it has a relatively large variance in age distribution, or if it is already in the Dividing stage. For a normal agent not being affected by any external factor, it stays in the Growing stage until reaching its full size (double its original volume). The length of the Growing stage should be the time needed for agent to double its volume. In the model I also set a typical growing time with a degree of variability (mean, standard deviation). However some drugs and the hypoxia condition may arrest agents at certain stages (see below). The arrested agents die when they reach the specified age.

The length of the Dividing stage is fixed and will not be affected by drug action or agent age. This means that if an agent gets into the Dividing state, it will definitely

split no matter whether it reaches the maximum age or not. The die stage is as an instantaneous process. Dead agents are moved out of group immediately. Although it is known that real agents take time to go through the apoptosis process, this stage is not the focus of my study so I introduce this simplification.

5-fluorouracil (5-FU) is drug used to treat a number of solid tumours. 5-FU competes with the natural substrate for the enzymes in DNA synthesis. The complex formed by 5-FU and enzyme is unable to perform the normal catalytic reaction which is crucial for DNA synthesis. In this way the cell affected by 5-FU stays in DNA synthesis phase of cell cycle and is therefore 'arrested' from normal cell cycle and unable to proliferate (David O. Morgan. 2007).

Oxygen is required for growth of tumour cells. Lack of oxygen supply, or hypoxia, is a common phenomenon in the inner layer of solid tumour (Drasdo and Hohme, 2005). Hypoxia also triggers several pathways such as HIFs (hypoxia inducible factors) and VEGF (vascular endothelial growth factor) (Shirinifard *et al*. 2009). The former promotes cell survival (Ortmann *et al*. 2014), and the latter induces vascular growth near tumour cells (Shirinifard *et al*. 2005). Also both simulations in (Drasdo and Hohme, 2005) and (Shirinifard *et al*. 2005) show a slowed cell cycle with hypoxia, and study in (Yijun Chen *et al*. 2009) shows a slowing of tumour cell proliferation with moderate hypoxia.

### 5.1.2 Population cell growth modeling

The group population growth of cells is another mechanism that needs to be added to model, Which I implement by introducing the ‘death rate’. The death rate represents the chance that an agent dies when it gets to its typical age. Although for single agent, it only controls the choice of death or division, the death rate may affect the emergent behaviour with a group of agents. As shown in Figure 5.2.1, by drawing the number of agents over time one can obtain a growth curve (see Section 2.3.1). In the model the death rate is represented by a real value between 0 and 1; 0 means 0% death rate, with which the agent will certainly divide; 1 means 100% death rate with 0% death rate the agent will certainly die.

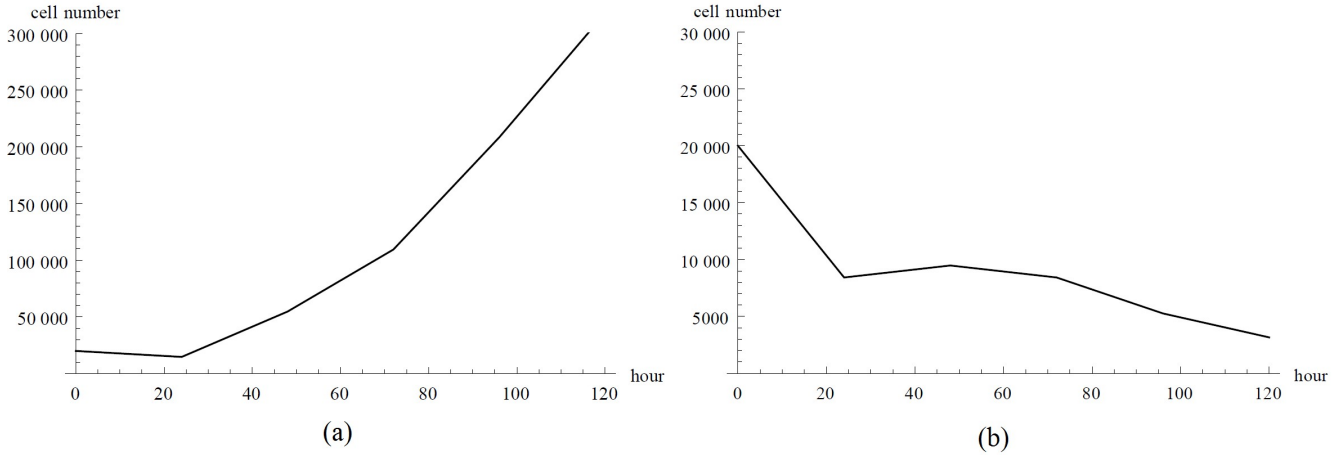



Figure 5.1.2 Typical growth curve with increase (a) and decrease (b) total number of agent (graph is drawn with experimental data of control group and combination group; to clearly show the trend of increase/decrease, the scales in (a) and (b) are different, see Section 5.5 for formal plot with same scale.)

I periodically counted the number of agents in all experiments and so can plot the population growth curves. By comparing experimental growth curves with simulation growth curves I can determine how well the model is calibrated with respect to population dynamics. The death rate may be affected by the environment and particular characteristics of the cell line. Since I grow cells of a single cell line in petri dishes with homogeneous media, I assume all the agents have the same death rate, while the value of death rate may vary over time due to the change of environment (temperature fluctuations, drug effect degradations, etc.).

In ideal conditions the space and nutrient levels for each cell are unlimited, so that the total number of cells has no limit either and obeys an exponential rule (Antonio Brú *et al*. 2003). The exponential rule can be written as equation as follows.

$$\frac{dN}{dt} = r_{Grow} N \tag{1}$$

Where N is number of agents, $\frac{dN}{dt}$ is change of number of agents, and $r_{Grow}$ is specific growth rate. If $r_{Grow}$ > 0, the number of agent increases; if $r_{Grow}$ = 0, the number of agents stays the same; if $r_{Grow}$ < 0, number of agent decreases. Integrating (154) I can get

$$\ln N + C_1 = r_{Grow} t \tag{2}$$

$$e^{r_{Grow} t - C_1} = N \tag{3}$$

In which $C_1$ is a constant. To determine its value, I assume at time $t = 0$, the number of agent is $N_0$. So that

$$\ln N_0 + C_1 = 0 \tag{4}$$

$$C_1 = -\ln N_0 \tag{5}$$

And equation (156) becomes

$$e^{r_{Grow}t+\ln N_0} = N_t \tag{6}$$

$$e^{r_{Grow}t} \cdot e^{\ln N_0} = N_t \tag{7}$$

$$N_0 e^{r_{Grow}t} = N_t \tag{8}$$

Now I let $t$ be the time that a group of agents doubles its population, and then I get $t = \ln\frac{2}{r_{Grow}} \approx \frac{0.693}{r_{Grow}}$. I call this $t$ as doubling time.

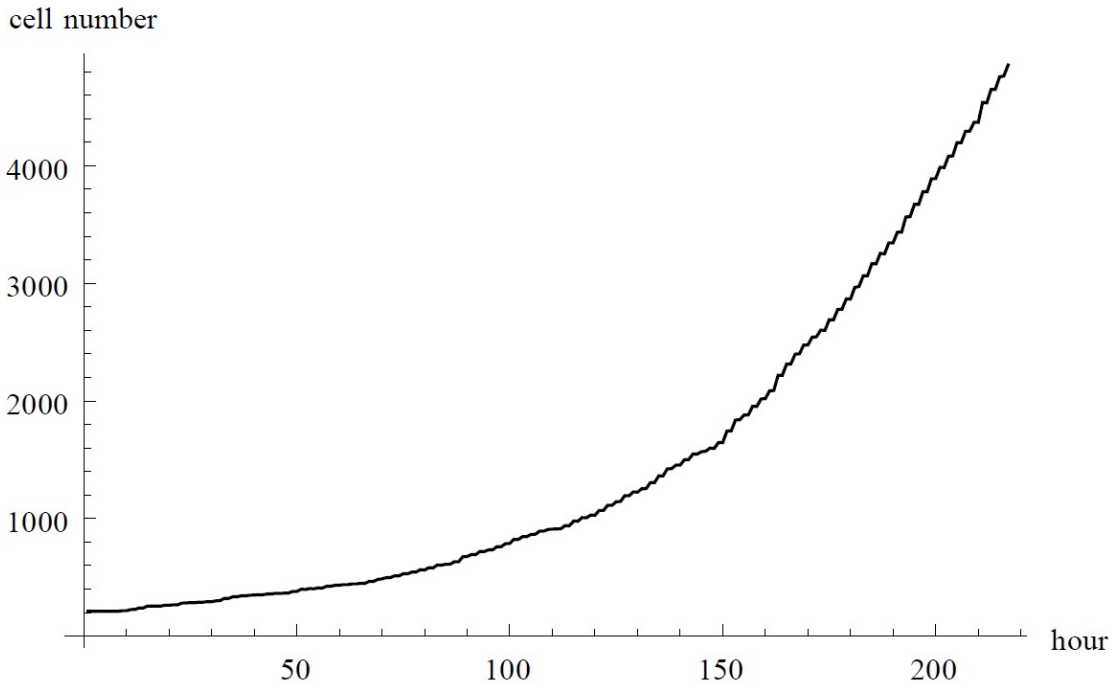



Figure 5.1.3 Typical exponential growth curve (graph is drawn with result of simulation with fixed $r_{Grow} > 0$)

In the real experiments, the size of petri dish is limited, and there will be not enough nutrients for an unlimited number of cells. Under this condition, the total number of cells is thus limited. I introduce a new variable called saturation density to represent

the limit that an environment may support, which is represented by $K$.

Then the previous equation becomes $\frac{dN}{dt} = r_{Grow} N\left(\frac{K-N}{K}\right)$. When $N > K$, $\left(\frac{K-N}{K}\right) < 0$, number of agent decreases; when $N = K$, $\left(\frac{K-N}{K}\right) = 0$, number of agent stays the same; when $N < K$, $\left(\frac{K-N}{K}\right) > 0$, number of agents increases.

Incorporating $\frac{dN}{dt} = r_{Grow} N\left(\frac{K-N}{K}\right)$ into model results in a growth curve shown as Figure 5.2.3.

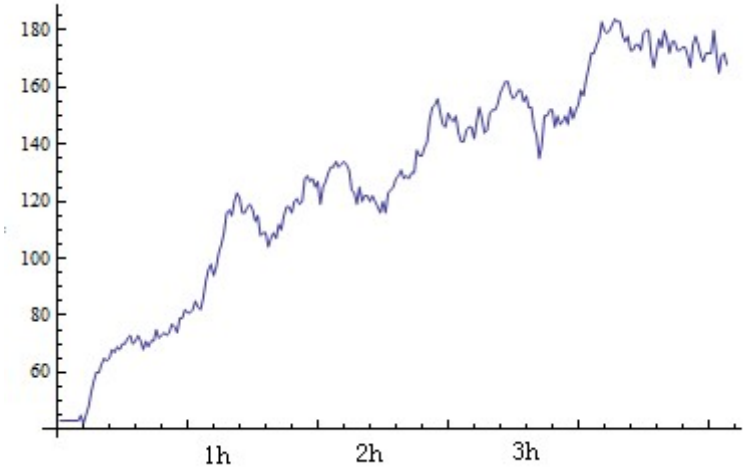



Figure 5.1.4 Typical limited growth curve with $K = 180$; x-axis is time, y-axis is number of cells; the zero value near the origin is produced by plotting software. After 3 hours, the curve waves around 180.

I assume that all the agents have the same death rate and all agents reach the end of cell cycle at exact same time. For $N$ agents with the same death rate $r_d$, there will be $N \cdot r_d$ agents die and $N \cdot (1 - r_d)$ agents dividing into two identical agents.

Therefore there will be $2N \cdot (1 - r_d)$ agents in the next cell cycle. I can substitute this into the previous equation $\frac{dN}{dt} = \mu N$, in which $\frac{dN}{dt}$ is the change of number of agent and equals $2N \cdot (1 - r_d) - N$. Thus I have $r_d = \frac{1 - r_{Grow}}{2}$. Because the growth rate is a population level behaviour, while the death rate is single-agent behaviour, this equation connects the two. However the initial simulation shows that the number of agents increases more slowly than in the experimental data.

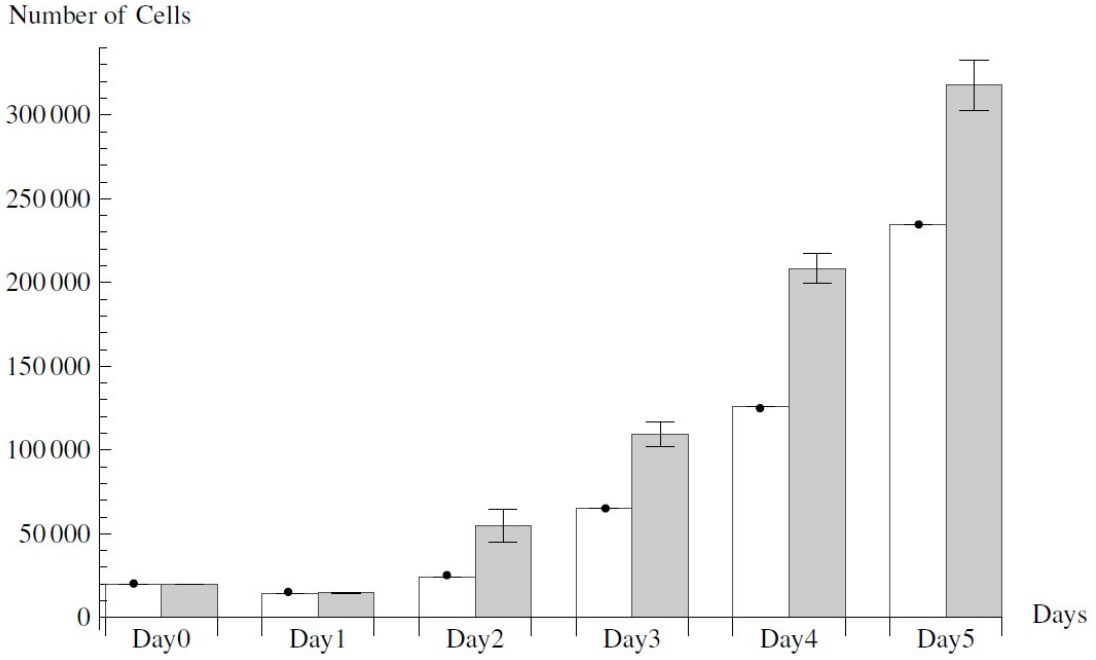



Figure 5.1.5 Population growth of simulation (white bars) and experiment (grey-scale bars). The number of cells in simulation increases slower than experiment

A possible explanation is that, in the experiment, not all the cells are in the exact same stage in the cell cycle. However by assuming the global growth rate, any differences in the cell cycle of each cell is ignored. Then in simulation I let the agents start with random cell cycle. In this way the growth rate in simulation is slightly smaller than experiment. So that by using this method I introduce inaccuracy. To account for this I have to avoid using doubling time. A new method is introduced in section 5.5 to

calibrate the model with experimental data in a better way.

## 5.2 Experiment design

### Experimental set up

Experiments are conducted using cell lines: cells grown for experimental use which have well-understood properties, such as the activation of particular oncogenes, or the ability to form structures such as spheroids. In the lab, growth experiments may be conducted on a Petri dish, in which case cells can grow into a structure (see Figure 4.2.3.1 for example). Our biological experts helped to choose HCT116 cell line, which is sensitive to drugs and easy to form structure on petri-dish or spheroid in gel.

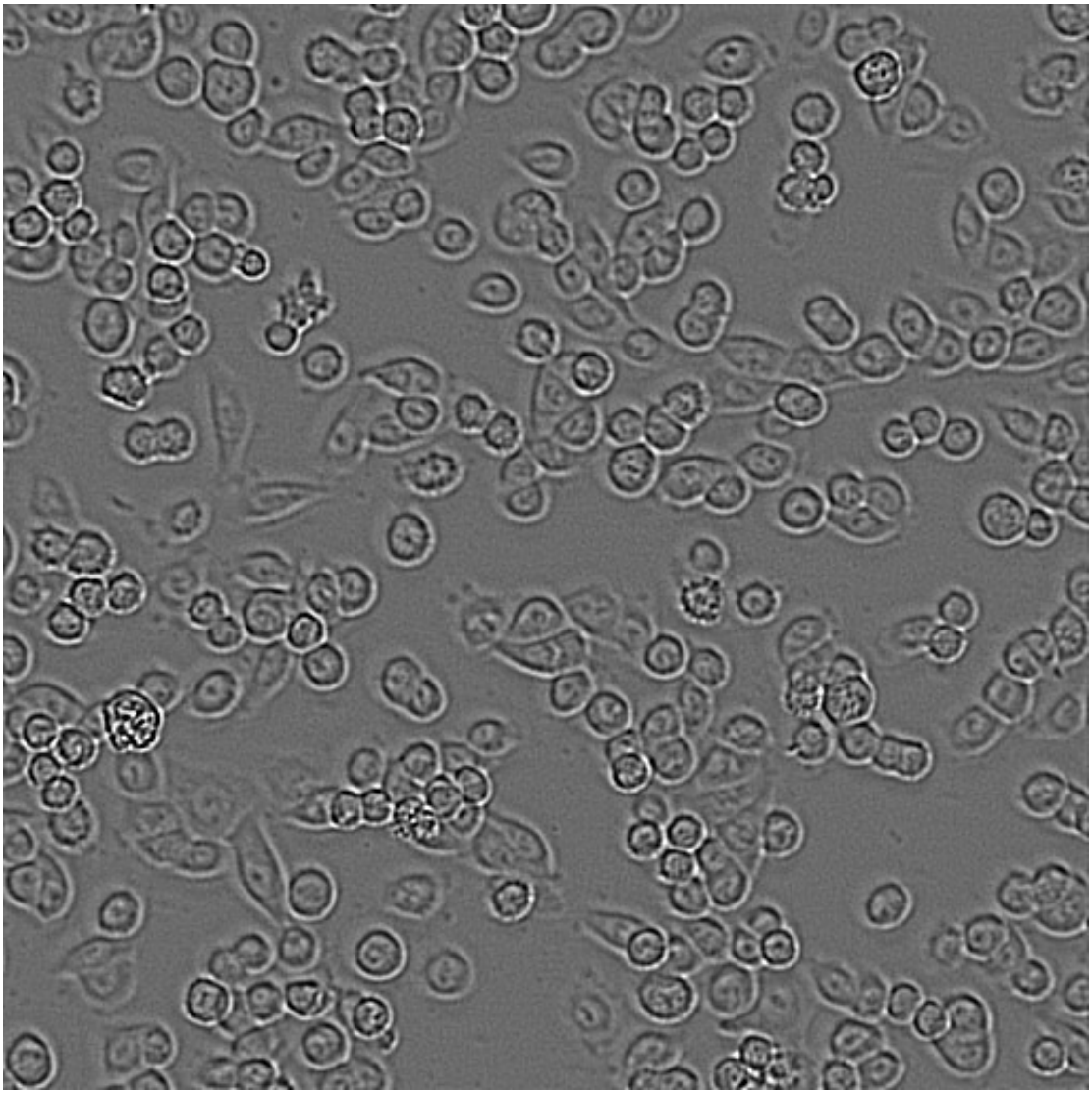


Figure 4.2.3.1 Cells growing on a glass plate (graph is taken from experiment carried out for Chapter 6)

I are primarily interested in the shape of structures formed by tumour cells, and how drug and environmental perturbations affect this structure. To clearly show the effect of environment elements, I have one experiment with no additional treatments to serve as a control, so that results from later experiments can be compared with results from the control experiment. I choose the therapeutic agent 5FU (Choudhary *et al*. 2012) to seek to understand how the structure is affected by drug. Note that the structure is linked with cell interaction, so by comparing normal structure and 5FU

structure I can analyse how this drug affects the interaction among cells. Similarly I want to see how a hypoxic environment affects the structure by comparing normal structure and hypoxic structure. And finally I will grow cells with 5FU in a hypoxic environment.

In addition, I wish to explore the differences in behaviour of different kinds of cancer cells. Our biological experts determined to use a normal cancer cell, or wild type cell; and a mutated cell line with p53 mutation which leads to a known degree of resistance to the drug and hypoxic environment. As discussed in previous paragraph, for each cell type, four experiments will be carried out, which are one control experiment, one with drug (5-FU), one in hypoxic environment, and one with drug in hypoxic environment. Two cell lines with four experiments each, thus there are eight experiments in total:

| No. | Cell line (HCT116) | | Environment | | |
|---|---|---|---|---|---|
| | p53 wild type | p53 mutated | No treatment | 5-FU | Hypoxia |
| 1 | √ | | √ | | |
| 2 | √ | | | √ | |
| 3 | √ | | | | √ |
| 4 | √ | | | √ | √ |
| 5 | | √ | √ | | |
| 6 | | √ | | √ | |
| 7 | | √ | | | √ |

| | | | | | |
|---|---|---|---|---|---|
| 8 | | √ | | √ | √ |

This set of eight experiments were carried by our biological experts and the result was analysed and used to calibrate the new model, see Section 5.2 and Chapter 6 for the analysis.

## Data to be collected

The experiments were designed, in collaboration with cell biologists, to ensure sufficient data to allow calibration of the physical parameters. Some parameters, such as the size of cell, can be directly obtained by measurement of a certain cell line. Some parameters, on the other hand, need to be measured during the course of the experiment or processed after experiment. The method of measurement depends on the type of data and its characteristic.

First of all, there are purely physical-related data. These data are used to calibrate the parameters of the physical model. I need data about cell size, shape and density to give these values to model cells, which as mentioned above, can be measured during experiments. I next need to know how many cells to put onto the substrate of how much area, so that the size of petri dish is also necessary. I also need length of experiment in order to control the simulation to run for a correct amount of time continually. The density and dynamic viscosity of fluid is also required, as they are related to value of constants to calculate resistance force and torque when cells are moving in experimental media.

Second, I need to add in the cell cycle model (cell growth and division etc.) and combine it with physical model, so I need data to calibrate growth-related parameters as well. I need the length of typical cell cycle of the cell line chosen, which corresponds to single cell growth curve, including a profile showing how cell size and shape can change as the cell cycle progresses. This can be determined by observing the cells over the course of the experiments. This information will need to be obtained by time-series imaging under the experimental conditions I wish to simulate. I also need the growth rate for the population per experiment, which requires to keep record of change of number of cells over time. With the population growth rate, I can calibrate the population growth curve. The growth rate is important because it controls the number of cells, which is essential for the correct mapping of cell distribution.

Finally, to analyse the cell interactions, I need to measure the position of the cells over time during each experiment. To get the cell position I need to take photographs of the cells over time, which will be processed by image analysis software. In this way I can obtain both number of cells and position of each cell at the same time.

## Methods to collect data during experiment

From the previous section it is known that the only data I can obtain during experiments are the change in number of cells over time, and the change in cell position over time. The experiments should contain several cell cycles so that the cell division and death can be observed. The number of cells should be counted

periodically, because it is not precise enough for the model to just match the starting and ending point. Note that I need non-destructive method to obtain this value; otherwise the experiment cannot be continued. Similarly, I need to image the growing cells at regular intervals as well: these photographs may need to be processed before to provide improved image sets. As mentioned before, I will use image analysis software to extract information from the images to obtain the position and number of cells.

I will then have directly compatible data from both experiments and simulation that can be analysed using a consistent approach. And although the structure grown in experiment and simulation may have a different appearance, I consider them match if their spatial distribution curve (as described in Chapter 6) is similar.

## 5.3 Experiment and result

The experiment is carried out following the design in Section 5.2 by our cancer cell biology expert. There are two stages to the experiment. In stage 1, both cell lines, HCT116 p53wt and p53mut, are used to determine the cytotoxicity of 5-FU (5-Fluorouracil) and hypoxia environment. In both stages, the hypoxia environment is mimicked by Cobalt chloride (CoCl2) and the cytotoxicity is represented by the death rate according to concentration of CoCl2. In stage 2, the 8 growth experiments are carried with the proper concentration of drugs determined in the first stage.

### Stage 1

The aim of stage 1 is to determine the proper concentration of the drugs by comparing the death rate of cell with various concentrations of 5-FU and CoCl2. The concentration cannot be too high or too low, with which the death rate of cell is too high or too low respectively. In stage 1 20,000 cells were seeded in every well of a 96-well tissue culture plate and left to grow for 18 hours. Then the media was changed and a series of concentrations of drugs were added. After 24 hours, the cells were subjected to Neutral red uptake based cell cytotoxicity assay to determine percentage cell death. The death rate of both cell lines with various concentrations is shown in Figure 5.3.1 and 5.3.2 respectively. The following two figures are produced by our cancer cell biology expert, Dr Hilal Khalil.

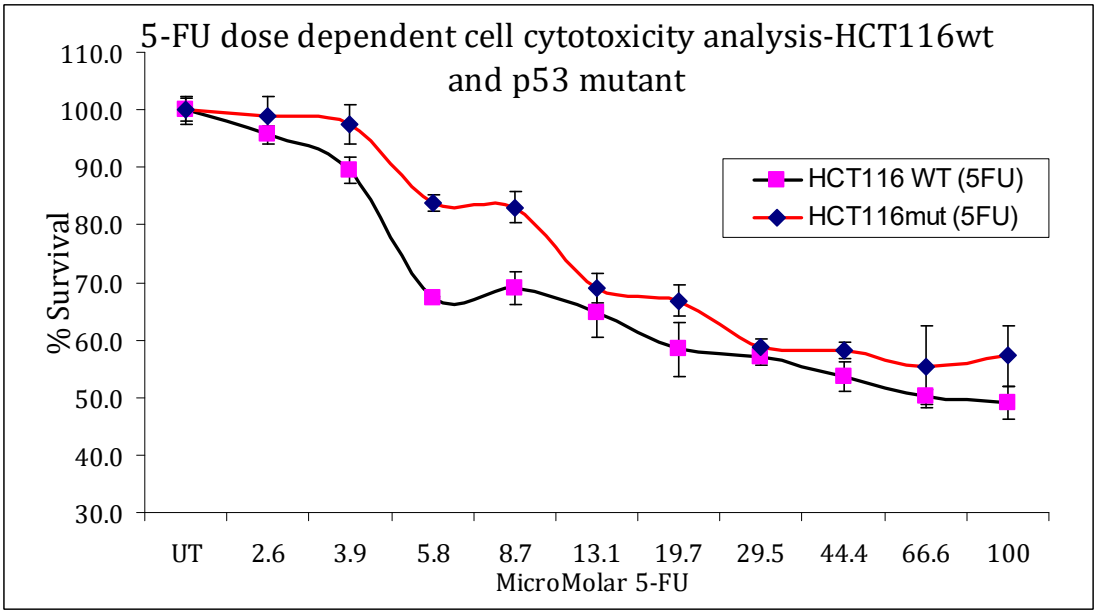



Figure 5.3.1 Survival rate of HCT116wt (black line with red dots) and p53 mutant (red line with black dots) cell lines under effect of various concentration of 5-FU. The x-axis is the concentration of 5-FU meansured by MicroMolar; and the y-axis is the survival rate of cell, which is measured by %.

From Figure 5.3.1, 'UT' means untreated, with which no 5-FU is added, thus the survival rate of both cell lines is 100%. Apart from the UT condition, there are 10

measurements for both cell lines. As the concentration of 5-FU increases, the survival of cell decreases. However the measurements with 2.6, 3.9 5.8 and 8.7 micromolar show that the survival rate of p31 mutant cell line does not differ very much from UT. The survival rates of two cell lines are too similar with concentration of 13.1, 29.5, 44.4 and 66.6 micromolar, which cannot show the difference of the two cell lines. The measurements with concentration of 100.0 micromolar shows a significant fall comparing with UT condition, and survival rates of two cell lines vary more than 10 10%. However the measurement of p53 mutant cell line contains a bigger error bar, which suggests bigger difference in the data of three countings. Therefore concentration of 19.7 micromolar is the best choise. In stage two, 20.0 micromolar is used.

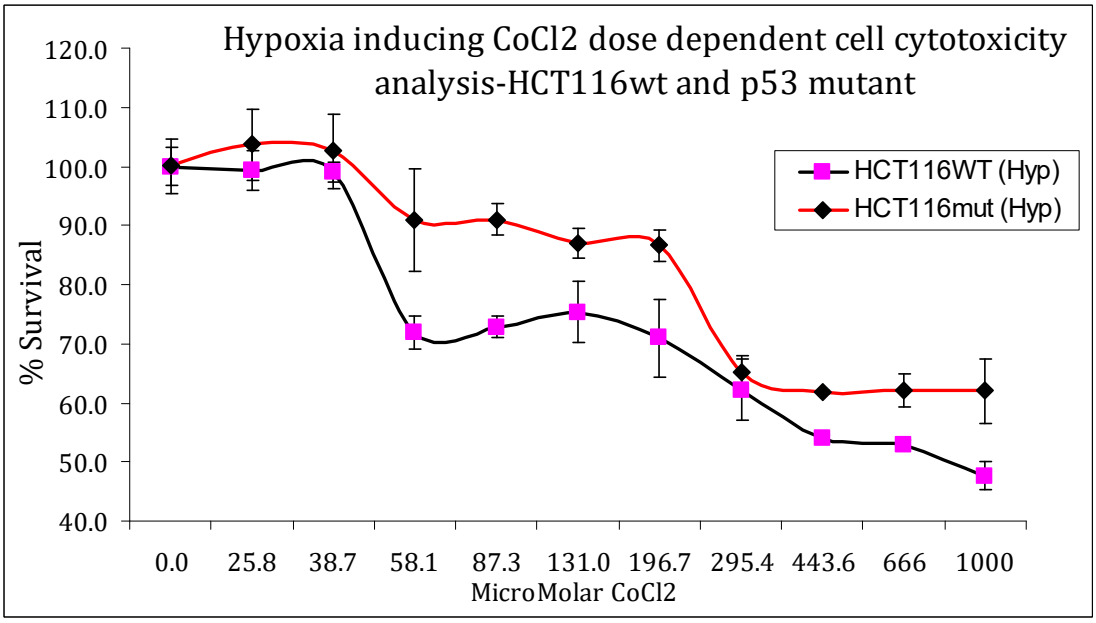



Figure 5.3.2 Survival rate of HCT116wt (black line with red dots) and p53 mutant cell lines (red line with black dots) under effect of various concentration of hypoxia drug. . The x-axis is the concentration of 5-FU meansured by MicroMolar; and the y-axis is the survival rate of cell, which is measured by %.

In the Figure 5.3.2, there are also 11 measurements for both cell lines, including the

UT condition (the survival with 0.0 MicroMolar). With concentrations of 25.8, 38.7, 58.1, 87.3, 131.0 and 196.7, the survival rate of p53 mutant cell line does not show significant fall comparing with untreated data. With concentration of 295.4 micromolar the survival rate of both cell lines are simiar, which cannot show the difference of the two cell lines against hypoxic environment. For the concentration of 1000 micromolar, the survival rate of p53 wild type (p53 wt) cell line is too low. Thus the concentration of 443.6 and 666 miromolar can provide better result. In stage two, 500 micromolar is used.

In addition, to show the survival rate of both cell lines with combination of 5-FU and hypoxia environment, both 20.0 micromolar of 5-FU and 500 micromolar of CoCl2 are added together.

## Stage 2

Stage 2 is to study the effect of drugs on growth rates. From stage 1, it is determined that the concentration of 5-FU and CoCl2 (for hypoxia) is as follows:

- 20microMolar 5-FU
- 500microMolar CoCl2 (Hypoxia)
- A combination of 20microMolar 5-FU + 500 microMolar CoCl2

In stage 2, both cell lines are used with the concentration of drugs listed above. Exactly 20,000 cells were seeded in each well in 96-well plate. This was considered day '0'. During next 5 days, cells in 3 out of 96 wells were sampled and counted each

day. By plotting the cell population of each day, I build the population growth curve of both cell lines, which are shown in Table 5.3.4 and 5.3.5. Note that 3 wells were counted for each day, all the number of cells in Table 5.3.4 and 5.3.5 are average value.

| | | HCT 116 wt | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Hrs | DAYS | UT | SD | 5-FU | SD | Hyp | SD | 5-FU/Hyp | |
| 0 | **0** | 20000 | 100 | 20000 | 100 | 20000 | 100 | 20000 | 100 |
| 24 | **1** | 14737 | 200 | 8421 | 546 | 13684 | 500 | 8421 | 415 |
| 48 | **2** | 54737 | 10000 | 24000 | 1125 | 11579 | 564 | 9474 | 564 |
| 72 | **3** | 109474 | 7400 | 63158 | 1456 | 23158 | 1087 | 8421 | 1087 |
| 96 | **4** | 208421 | 8900 | 77895 | 1147 | 13684 | 1040 | 5263 | 569 |
| 120 | **5** | 317895 | 15020 | 89474 | 2067 | 10526 | 789 | 3158 | 487 |

Table 5.3.4 The experimental population growth curves of HCT116 p53wt.

| | | HCT116 mut | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Hrs | DAYS | UT | SD | 5-FU | SD | Hyp | SD | 5-FU/Hyp | SD |
| 0 | **0** | 20000 | 1850 | 20000 | 10500 | 20000 | 1050 | 20000 | 1050 |
| 24 | **1** | 17500 | 9090 | 15000 | 18000 | 20000 | 1200 | 12000 | 1213 |
| 48 | **2** | 55000 | 16054 | 30000 | 10405 | 27500 | 1547 | 12500 | 604 |
| 72 | **3** | 202500 | 28014 | 80000 | 2809 | 22500 | 147 | 12500 | 478 |
| 96 | **4** | 477500 | 25032 | 192500 | 25014 | 20000 | 287 | 15000 | 1009 |
| 120 | **5** | 812500 | 37000 | 195000 | 29874 | 20000 | 201 | 10000 | 743 |

Table 5.3.5 The experimental population growth curves of HCT116 p53mut.

The time-lapse images are taken on the same well every 10 minutes for 18 hours. The sample photograph is shown in Figure 5.3.6. The size of each time-lapse image is $512\times512$ pixels, and by comparing the relative size of whole image and each cell I can therefore estimate that the diameter of cell is about 10 pixels. From the series of photographs it is also able to estimate the length of typical cell cycle.

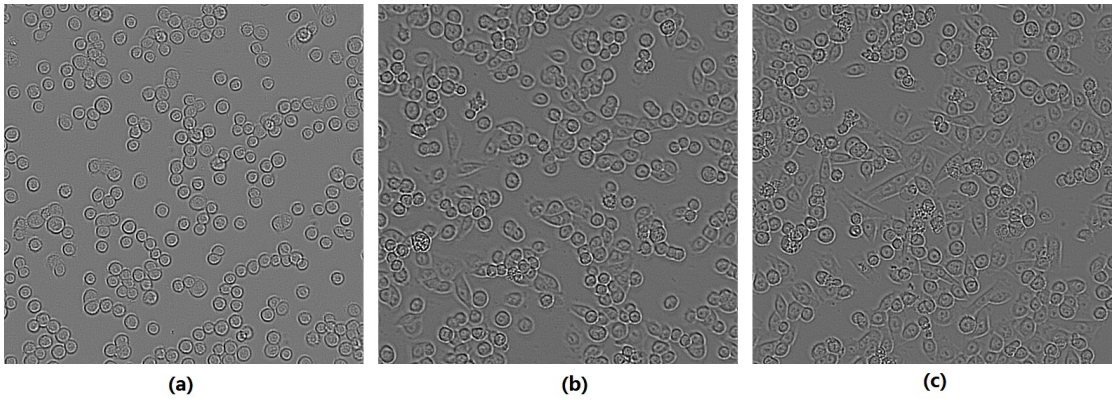



Figure 5.3.6 Time lapse images taken from growth experiment of HCT116 p53mut cell line with effect of 5FU: the left image is the start point of experiment; the middle and right images are taken of the same area with interval of 10 hours.

## 5.4 Model parameterisation

In the simulation, there are some other parameters that relates to experimental data. These parameters are verified as follows.

### Cell size and shape

Compared to endothelial cells, the cell line used for the cancer cell interaction study has a different size and shape. Due to the importance of cell size and shape to the physical interaction, I need to modify the relevant values in the model.

The shape and size of cell can be estimated from experiment photographs using $r$ to

represent the sphere radius, so that its volume is $V_0 = \frac{4}{3}\pi r^3$.

When the cell is attached to the petri dish, its shape and size may vary. I need to find a typical size and fit it to an ellipsoidal shape. Then I can use it as agent shape/size in simulation. To define the shape of an ellipsoid, I need to know the length of its three semi-axis. By measuring the cells in experiment photographs I can estimate the length of two semi-axis which are horizontal to the petri dish, and with the volume of un-attached agent I can calculate the last semi-axis by ellipsoid volume equation $V_e = \frac{4}{3}\pi a_1 a_2 a_3$, where $a_1$, $a_2$ and $a_3$ are semi-axes of ellipsoid and $a_1 < a_2 < a_3$ (in the model I have the rule that $a_1 < a_2 < a_3$, see section 3.1 for detail).

The length of three semi-axes is determined by estimating their ratio while keeping the volume of agent as $\frac{4}{3}\pi r^3$.

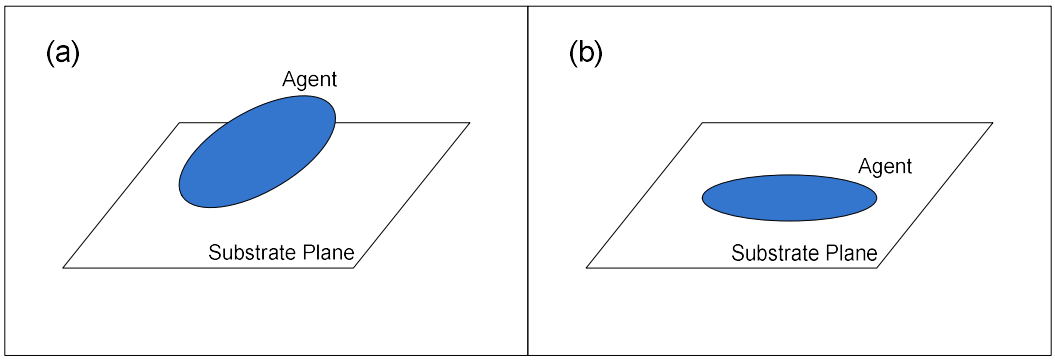



Figure 5.4.1 (a) Un-attached agent; (b) Attached agent becomes more flat.

In the experiment I can observe that cells attach to the bottom of petri-dish and the shape is generally flat. The semi-axis that is vertical to the petri dish should be much

shorter than horizontal semi-axes. In addition, to afford the potential polarity of agents the two semi-axes parallel to the petri dish plane should not be equal to each other. However, from time-lapse images it is known that the ratio of these two semi-axes should not be too large. A value between 1:1 and 1:2 may be appropriate. I let semi-axis $a_1 = \frac{1}{4}r$ , $a_2 = \frac{\sqrt{2}}{2}r$ , $a_3 = 2\sqrt{2}r$ ; so that the volume of agent remains $\frac{4}{3}\pi r^3$ . From the photographs of experiments, I can determine the diameter of a typical HCT-116 cell is 10 $\mu m$ immediately after division. Thus $r = 5\mu m$ . Since the unit length $\Delta x = 1.25\mu m$ , measured by the unit length, $a_1 = \Delta x$ , $a_2 = 2\sqrt{2}\Delta x$ , $a_3 = 8\sqrt{2}\Delta x$ .

## Cell age

By observation the time-lapse images, I determine that cells in the experiment divide after 12 hours. During this time the volume of the cell is doubled. Therefore I take the value of 12 hours as the typical cell cycle in experiment. In the real experiment, the growth of cells may vary, so that I set 12 hours as the expected value and add 10% variance. Also in the real experiment each cell may start growing from a different position in its cell cycle. To simulate this, I let each simulation agent start from a random position for its age as well. Figure 5.4.2 shows the sample growth curves with and without both the cell age variance and random agent starting up position, in which the curve with cell age variance and random agent starting up position is smooth.

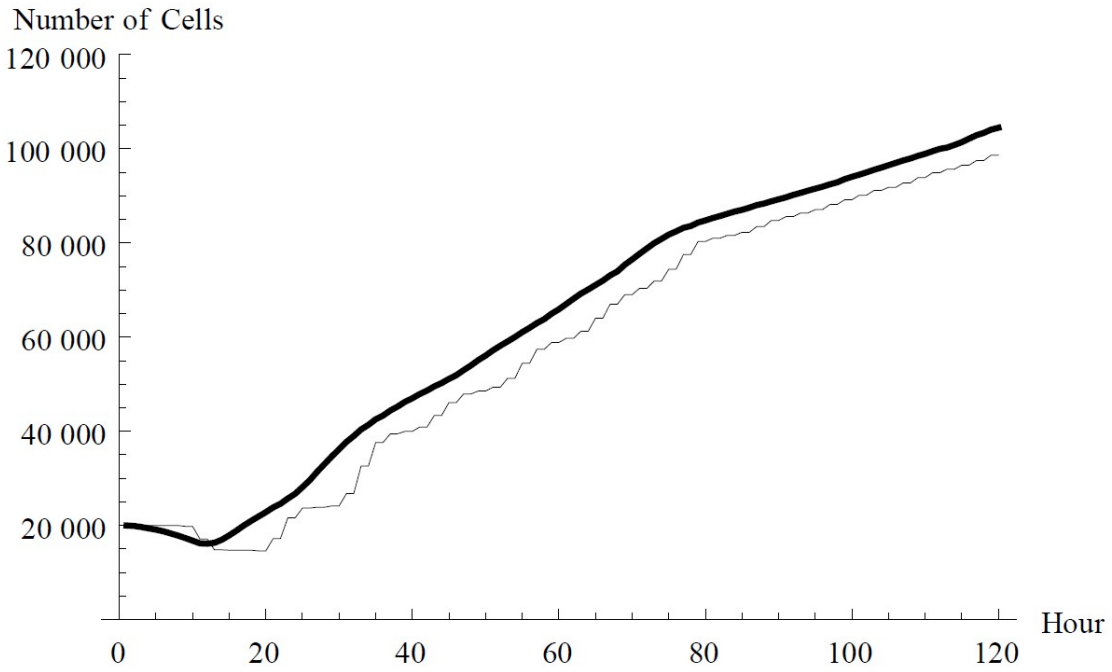



Figure 5.4.2 Stepped curve (thin line) and smooth curve (thick line) (graph is drawn with dynamic growth rate $r_{Grow}$. Note that the smooth curve has higher value of number of cells, because with random agent starting up position in cell cycle, some agents divide earlier and enters high growth rate period earlier; thus more agents are produced.)

I let simulated agents complete one modelled lifespan in 100 time steps. Thus 100 time steps equal 12 hours in the real experiment. As in the experiment the total number of cells is counted every 24 hours, and the total number of simulation agents is extracted every 200 time steps. Note that I output to a file the number of cell data every 10 time steps, the interval of measurement in experiment equals same simulation time to produce 20 outputs. Therefore if curves from the experiment and simulation are drawn on same coordinate system the data point for 'Day 1' of experiment should be plotted on x-coord=20.

A setting file stores all the growth information for each cell line which is read and set to agents at simulation start up. The following 9 settings form a typical set of

information for one simulation:

(1) celltype = 1

There are several sets of built-in shape of ellipsoid in the simulation, stored in a two-dimonsion array. The value of ‘celltype’ is used to access the array to find the corresponding size of semi-axice of cell shape;

(2) numberofphase = 5

Because there are six measurements of total number of cells in our experiments, and it is assumed that in each interval the growth rate of cell is constant. There are five intervals in six measurements, thus each interval is considered as a phase of growing. This setting works together with setting (5)-(9). The value of ‘numberofphase’ is read then the code knows how many phases to read, so that it can prepare space to contain growth rate of each phase;

(3) MaxAge=100

The “MaxAge” is the expected value of typical cell cycle length. As previously it is assumed that 100 time steps in simulation equal 12 hours in the real experiment, here the value 100 is assigned to typical cell cycle length;

(4) AgeVariance=10

In the previous section, the length of cell cycle is assumed to obey normal distribution. The “AgeVariance” is the variance of the distribution. Using ‘MaxAge’ and ‘AgeVariance’, each agent in simulation can has different length of cell cycle;

(5) 20, 0.42920

(6) 40, 0.96362

(7) 60, 0.70711

(8) 80, 0.68990

(9) 100, 0.61751

Setting (5)-(9) are working together with setting (2) to control the probability of cell division in the simulation (the definition of probability of cell division is at the beginning of Section 5.5). The setting (2) regulates how many phase are in simulation, and settings (5)-(9) regulate value of probability of cell division in each phase. Each one of these five settings contains two parts: the first part is the end point of current phase, and the second part is the probability of cell division in the current phase.

Take the setting (5) as an example. The value '0.42920' means the value of probability of cell division should be 0.42920; and the value '20' means this value is valid until the program produces 20 outputs.

Note that the integer part (20, 40, 60, 80, 100) has different meaning from 'MaxAge' and 'AgeVariance'. The integer value for 'MaxAge' and 'AgeVariance' means the number of time steps, while the integer value for probability of cell division means the number of output file. As there is one output file every 10 time steps, the integer 20 here actually means 20x10=200 time steps, which equals 24 hours, which fits length of interval between day0 and day1 in the experiment.

When the program produces 20 outputs, the setting (5) is useless. Thus from that time point ($200^{th}$ time step, or $24^{th}$ hour) the setting (6) takes effect. Similarly the setting (6) is valid until there are 40 output files, in other words the setting (6) is valid between the $24^{th}$ hour to $48^{th}$ hour.

In this way these five setting define the value of probability of cell division for the simulation using data from five intervals of the *in vitro* experiments. The method to determine the value of probability of cell division from exeperiment is discussed in section 5.6.

In the same manner of setting (1)-(9), several sets of data can be defined to represent a unique cell type, and by assigning different cell type to agents, agents with different behaviour can be added in one single simulation.

### Time frame

In the *in vitro* experiment, the time-lapse images are taken every 10 minutes. The time frame of model is modified to this value.

### Dynamic viscosity of media

According to the documentation provided by the media manufacturer (http://www.lifetechnologies.com/uk/en/home/life-science/cell-culture/mammalian-cell-culture/classical-media/dmem.html), the media can be physically considered as water so that the value of dynamic viscosity does not need to be modified.

## 5.5 Population growth curve fitting

From the experiment data (Table 5.3.4) the population growth rate during each interval can be calculated. Note that the population of cell is counted every 24 hours in the experiment and, because of a lack of data between each measurement point; I

assume the population growth rate is constant throughout each interval, although the growth rate during each interval may of course vary. The population growth rates for the four experiments are shown in Table 5.5.1.

| HCT-116wt | UT | 5-FU | Hypoxia | 5-FU + Hypoxia |
|---|---|---|---|---|
| 0-24hr | -0.305 | -0.485 | -0.380 | -0.865 |
| 24hr-48hr | 1.312 | 1.047 | -0.167 | 0.118 |
| 48hr-72hr | 0.693 | 0.968 | 0.693 | -0.118 |
| 72hr -96hr | 0.644 | 0.210 | -0.526 | -0.470 |
| 96hr -120hr | 0.422 | 0.139 | -0.262 | -0.511 |

Table 5.5.1 Population growth rate of 4 experiments with HCT wild type cell line

The population growth rate cannot be directly used in the simulation, because I use an agent-based model and what is really needed is the chance that the agent divides at the end of its cell cycle.

I introduce a new term, “probability of cell division”, to the simulation to control the chance that agent successfully divides. The probability of cell division is a real number between 0.0 and 1.0, which represents the chance that one single agent divides at the end of its cycle: the maximum value means the agent will certainly divide; the minimum value means the agent will certainly die. The higher the value, the more likely it is that the agent will divide. For a group of agents, since they share the same population growth rate value, they also share the same value for the probability of cell division. From the definition of probability of cell division, it is

readily seen that if its value is higher than 0.5, the population of agents increases. If the value of the probability of cell division reaches 1.0, the population of agents satisfies exponential increase, which is described in section 5.2; if the value of the probability of cell division reaches 0.0, all the agents die.

Note that the probability of cell division has a different meaning to population growth rate: the probability of cell division takes effect when cell goes to the end of the cell cycle and so affects each single cell cycle; in contrast the population growth rate is a measurement from experiment and can be the result of several cell cycles. The value derived from the experiment is the population growth rate, but the value to be used in the model is the probability of cell division. The probability of cell division is a real number between 0.0 and 1.0, while the population growth rate does not have limit, specially for one single cell cycle, the population growth rate is in the range $[-1,1]$. Thus a conversion is required for the population growth rate and probability of cell division during single cell cycle.

Let $r_{Grow}$ equal the population growth rate, $p_{Div}$ equals the probability of cell division. As discussed in section 5.2, the population growth rate is calculated without using doubling time. It is known that the length of typical cell cycle is 12 hours. Therefore in each 24 hour-interval there are 2 cell cycles. As discussed previously, I start with $N_0$ cells; let $x$ represent the percentage of cells that divide, obviously after one cell cycle there will be $2x \cdot N_0$ cells left, because $x \cdot N_0$ cells divide to

$2x \cdot N_0$ cells and the rest of cells die. After one more cell cycle there will be $4x^2 \cdot N_0$ cells left. Using $N_1$ to represent the number of cells after 24 hours (2 cell cycles), then the equation can be written as follows

$$N_0 \cdot 4x^2 = N_1 \tag{1}$$

By solving it I have

$$x = \sqrt{\frac{N_1}{4 \cdot N_0}} \tag{2}$$

Therefore for each interval, with both measurements of number of cells before and after the interval I can calculate the percentage of cells that divide to two daughter cells, which I use on each agent as its chance to divide.

Also consider $N_1$ as $N_t$ in the first equation, the probability of cell division $p_{Div}$ satisfies $N_0 \cdot 4x^2 = N_t$, substitute it to the equation (161), I have

$$r_{Grow} = \frac{2}{t} \ln 2 p_{Div} \tag{3}$$

In which $r_{Grow}$ is the population growth rate, $p_{Div}$ is the probability of cell division for each agent, and $t$ equals 24 hours. In this way I build the connection between the behaviour of each agent and the behaviour of the population. However this connection is based the fact that the measurement interval in the experimental result contains 2 full cell cycles, and thus is experimentally specified. For any other experiments with different cell lines and measurement intervals the connection should vary. The population growth rate $r_{Grow}$ and probability of cell division $p_{Div}$ (according to the

equation (164)) of control population are presented in Table 5.5.2. Note that each interval of measurement in experiment contains two cell cycles, thus the population growth rate is not limited to $[-1,1]$.

| | 0-24hr | 24hr-48hr | 48hr-72hr | 72hr -96hr | 96hr -120hr |
|---|---|---|---|---|---|
| Population growth rate $r_{Grow}$ | -0.305 | 1.312 | 0.693 | 0.644 | 0.422 |
| probability of cell division $p_{Div}$ | 0.42920 | 0.96362 | 0.70711 | 0.68990 | 0.61751 |

Table 5.5.2 Population growth rate and probability of cell division of control group experiment

In the simulation each agent is assigned the same probability of cell division $p_{Div}$. From the definition of probability of cell division $p_{Div}$, it is obvious that if $p_{Div}$ is higher than 0.5, the total number of cell increases; if $p_{Div}$ is lower than 0.5 the total number of cell decreases.

The next step is to map the growth curve of each experiment to the simulation. I map the control population first, using the probability of cell division calculated from experiment data. Then instead of using experiment data to map drug and hypoxia affected growth, I analyse the effect of these factors on the cell and introduce the mechanism into the model that changes the model behaviour in order to reach the

observed population growth curve. Finally, I then test my mechanism assumptions against independent data, specifically the population growth curve for the combined drug and hypoxia condition.

### 5.5.1 Model calibration with Control population

I apply the probability of cell division in Table 5.5.2 to the simulation and run for 10 times, simulation result is shown in Table 5.5.3.

| | Number of cells | | | | | |
|---|---|---|---|---|---|---|
| Simulation No. | 0 | 24 hr | 48 hr | 72 hr | 96 hr | 120 hr |
| 1 | 20001 | 29831 | 79705 | 149109 | 254143 | 374142 |
| 2 | 20002 | 29728 | 79430 | 146807 | 249908 | 368676 |
| 3 | 20003 | 29650 | 79155 | 147719 | 251822 | 371681 |
| 4 | 20002 | 29990 | 80324 | 150097 | 255907 | 377912 |
| 5 | 20003 | 29686 | 79657 | 147657 | 250946 | 370312 |
| 6 | 20002 | 29882 | 80090 | 148767 | 253311 | 373235 |
| 7 | 20003 | 29821 | 79729 | 147904 | 251395 | 370854 |
| 8 | 20001 | 29627 | 79426 | 147428 | 251350 | 370295 |
| 9 | 20003 | 29850 | 80368 | 148933 | 252525 | 373784 |
| 10 | 20001 | 29804 | 79564 | 147848 | 250390 | 369295 |
| Average | 20002.1 | 29786.9 | 79744.8 | 148226.9 | 252169.7 | 372018.6 |
| SD | 0.8756 | 113.3504 | 399.1438 | 974.7133 | 1839.7811 | 2784.5162 |

Table 5.5.3 Cell number in 10 simulations of control group

I draw the data as barcharts with errors bar together with experiment data. As Figure 5.5.1.1 shows, before calibration the simulation data does not match the experimental result.

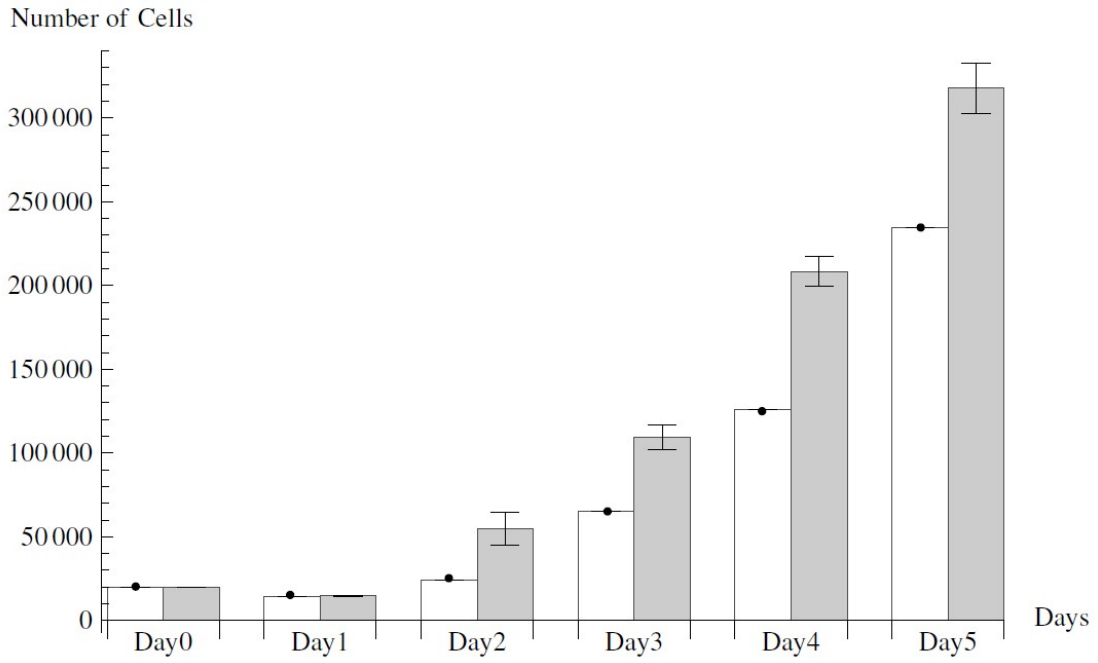



Figure 5.5.1.1 Control population growth curve without calibration, the white bars are simulation data, the grey bars are experimental data. The simulation ran once, so there is no error bar. I can see that the simulation population increases more slowly than the experimental population.

Figure 5.5.1.2 shows an improved result following calibration. Comparing to experiment data, the simulation result is always slightly higher. The reason for that is, in the calculation I assume all the agents have exact same cell cycle and start from time 0 but in the model agents have variation of 10% of their cell cycles and in simulation they start from random position in cell cycle. Thus some of the agents process to division stage earlier than theoretical situation, which lead to more

daughter cells. However the simulation result is sufficient close to the experimental data and can be accepted as baseline of other simulations.

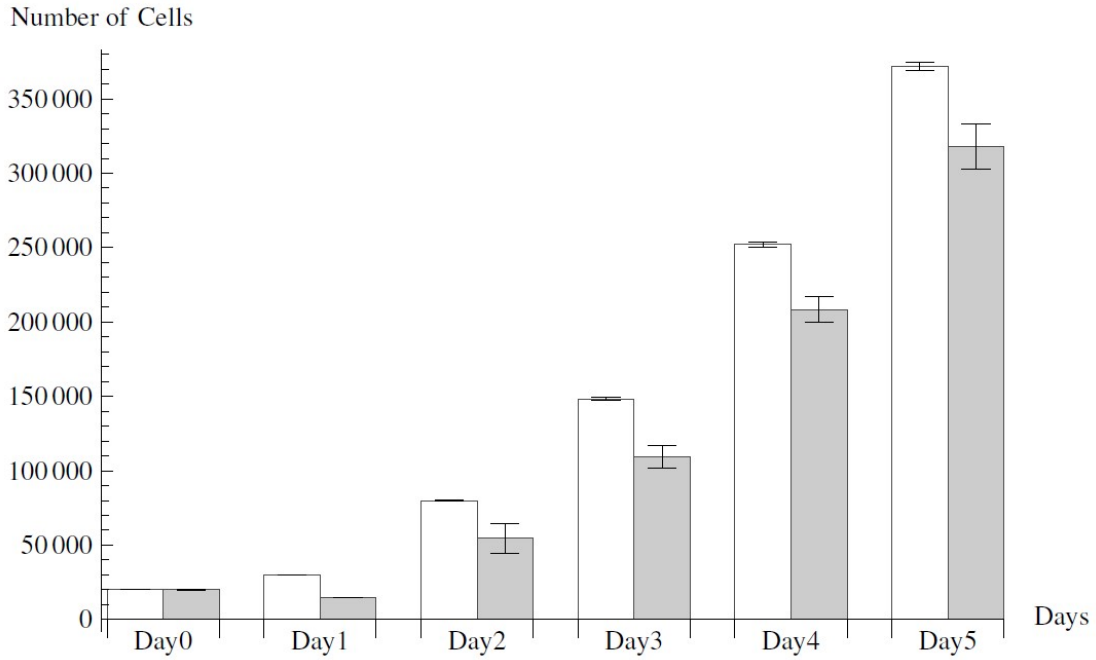



Figure 5.5.1.2 Control population growth curve with calibration. The white bars are average value of 10 simulations, the grey bars are experimental data. I can see that although the simulations do not match with experimental data, by adding variation to cell cycles and letting cells start from random position in cell cycle, the simulation result is closer with experimental data.

## 5.5.2 5-FU and the population growth curve

As discussed in Section 5.1, 5-FU disrupts DNA synthesis in tumour cells and therefore arrests tumour cells in DNA synthesis stage in cell cycle (Choudhary, *et al*, 2012). Thus tumour cells cannot complete cell cycle and finally die. Thus the mechanism of 5-FU can be considered as a decrease of cell proliferation. In my model, the decrease of cell proliferation is presented by reduction of chance of cell division. Three strategies have been considered in modelling how 5-FU affected cell growth. The first is to calculate the probability of cell division from the experimental data of

the 5-FU affected group, then use the value directly, as per the control population growth curve. The first strategy does not reflect mechanism of 5-FU and theoretically in this way I can map any population growth curve to model. The shortcoming is that it does not help to understand what is happening in terms of the cell responses to drug action. The second strategy is to amend the probability of cell division derived for the control condition in a way that reflects the 5-FU mechanism such that this probability works with the parameter set to fit to the control population growth curve to generate the simulation growth curve for the 5-FU condition. The third strategy is to let the probability of cell division change dynamically with the concentration of 5-FU. This strategy helps to uncover the connection between cell growth and 5-FU concentration, but it cannot help to show the difference between the controlled condition and the 5-FU condition. Therefore I choose the second strategy for all the following simulations, in which I will try to introduce a method to amend the control data to reflect the effect of drug (and then hypoxia).

To simulate the effect of 5-FU, a new state is added: Arrested, to the state machine previously shown in Figure 5.1.1. The arrow pointing from the Growing state to the Arrested state means that an agent can be affected by 5-FU machenism while it is growing. Once an agent is in the Arrested state it cannot turn back to the Growing state or Idle state, and that is why this new state is called ‘Arrested’. An agent dies when it gets to its cell age in the Arrested state. Thus the effect of 5-FU can be considered as an extra rate of death.

In the simulation, instead of changing the probability of cell division directly, those cells that decide to divide are given a chance to enter the Arrested state in the fate-determination phase. All the arrested cells will not divide; instead, they die after certain time. In this way the Arrested state works as an extra chance of death.

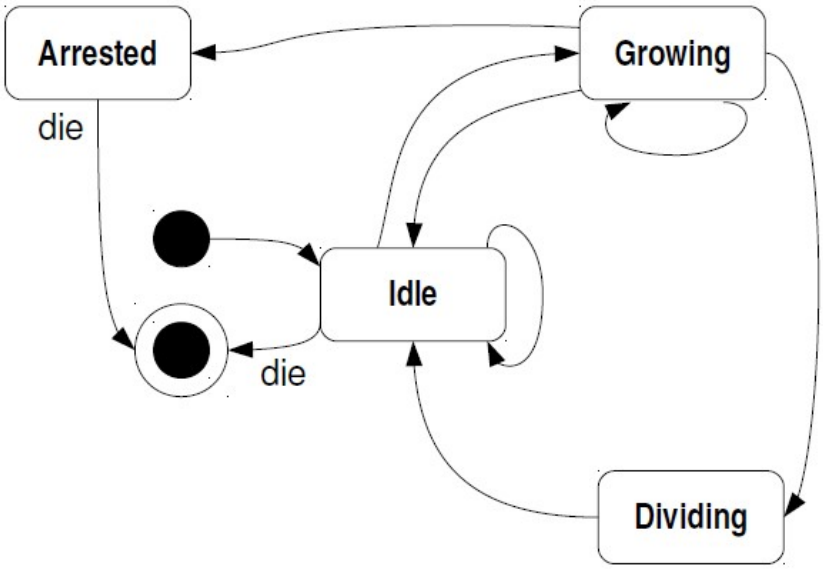



Figure 5.5.2.1 State of agent as state machine with mechanism of 5-FU modelled as 'Arrested' state

In this way there are several values to determine: how much a cell has already grown before it is arrested; the percentage of cells that go into 5-FU status. In these variables, the percentage may change over time, as it depends on concentration of 5-FU which is changing over time as well, but for simplicity I consider a constant arrest rate at the moment.

If the arresting chance is too high, the total number of cell will decrease, and if it is too low the growth curve will look like the controlled growth curve. After trying several values, I get the best simulation result when it is set to 5%, as shown in Figure

5.5.2.2.

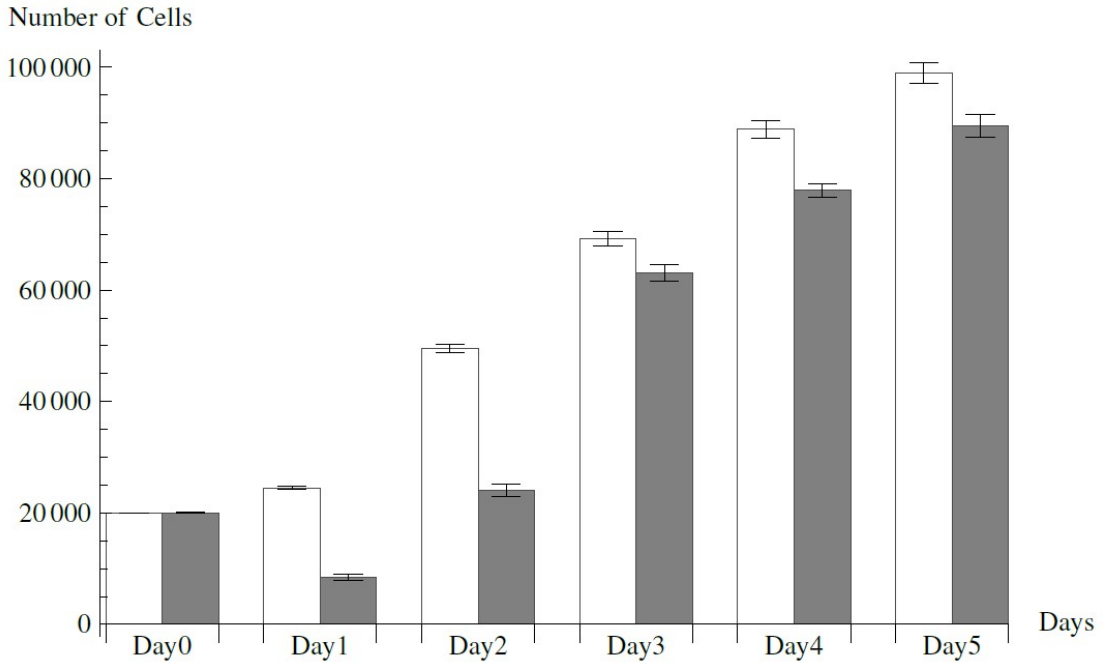



Figure 5.5.2.2 Simulation with 5% arrested agents compared with experimental result; the simulation bars are in white and the experiment bars are in grey. The value of simulation data is average of 10 runs, and the experimental data is from 3 counts. The simulation and experimental data is distanced in Day1 and Day2 but become closer in Day3, 4 and 5. Although the simulation and experimental data do not exactly match, they are efficient to show the added mechanism of 5-FU to the model.

## 5.5.3 Fitting to Hypoxia population

In a similar manner to the approach of mapping the effect of 5-FU on cell growth, I first analyse the biological effect of hypoxia on cell growth and then introduce a plausible mechanism to the model. Note that as I discussed in section 5.5, my cell cycle model is a simplified version of the underlying biology. In my model I ignore the inter-cell biological reaction and focus on cell behaviours that can be observed from outside, such as change of cell volume, cell division and cell apoptosis.

Then to introduce the mechanism of hypoxia, I first analyse its effect on cell growth,

including whether it affects the length of cell cycle and probability of cell division. The first strategy used is to introduce an upper limit number of cells that may survive in a single petri-dish. Considering the petri-dish as a closed system, the oxygen and nutrient supply is limited, so that the cells planted in petri-dish cannot grow forever. This upper limit of number of cells is referred as saturation density. Not only is the total number of cells in single petri-dish affected by the saturation density; the rate that the number of cells change overtime, i.e. population growth rate, is affected at the same time, as described in section 5.2.

The problem is that the effect of saturation density $K$ is only significant when number of cell approaches $K$, while in the experiment the hypoxia acts on cells at any density and thus has a uniform effect during whole experimental process. In other words, the saturation density $K$ has a different biological mechanism of hypoxia and is not a suitable representation to be introduced in simulation.

As described in Section 5.1, on cell level, hypoxia causes a slowed cell cycle; on tissue level, hypoxia reduces the growth of tumour (Yijun Chen, *et al*, 2009). Thus to introduce the effect of hypoxia, I need to reduce the probability of cell division, which is the second strategy. I assume that cells are affected equally during the whole experiment; therefore I have to amend the probability of cell division for all 5 intervals in the same way. As mentioned at the beginning of section 5.5, an increase in the cell population requires the probability of cell division to be higher than 0.5, and a

decrease of cell population requires it to be lower than 0.5. From Figure 5.5.3.1 it is known that the probability of cell division from 0 to 48 hours should be lower than 0.5, then higher than 0.5 from 48 to 72 hours, and again lower than 0.5 after that, as Table 5.5.4 shows.

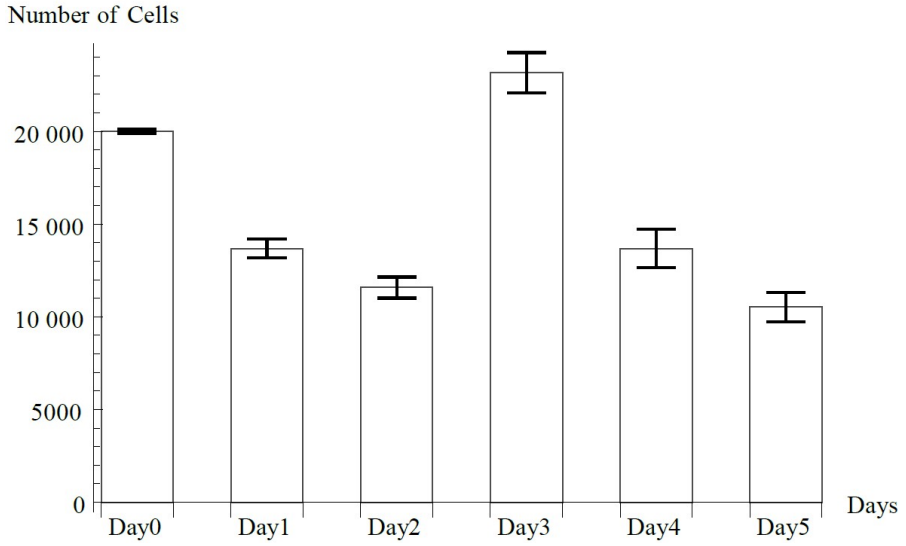



Figure 5.5.3.1 Growth curve of hypoxia (experimental), which is generated with 3 counts.

| | 0-24hr | 24hr-48hr | 48hr-72hr | 72hr -96hr | 96hr -120hr |
|---|---|---|---|---|---|
| Control condition probability of cell division | 0.42920 | 0.96362 | 0.70711 | 0.68990 | 0.61751 |
| Expected hypoxia probability of cell division | <0.5 | <0.5 | >0.5 | <0.5 | <0.5 |
| Method to amend control condition parameters | Reduce no more than 0.42920 | Reduce at least 0.46362 | Reduce no more than 0.20711 | Reduce at least 0.18990 | Reduce at least 0.11751 |

Table 5.5.4 Probability of cell division in experiment under control condition and method to amend the probability of cell division to fit hypoxia group

Table 5.5.4 also shows the method to amend the control condition parameters that may produce a well fitting growth curve. Even without recourse to simulation it is clear from Table 5.5.4 that the requirement between 24 and 48 hours cannot be satisfied as it will make the probability of cell division between 0-24 hr lower than zero and 48-72 hr lower than 0.5.

For the other 4 intervals, there is a value that satisfies all requirements: the original probability of cell division for 48-72 hour is 0.70711, and to keep it more than 0.5 it cannot be reduced by more than 0.20711; for 72-96 hour and 96-120 hour the probabilities are 0.6899 and 0.61751 respectively: to decrease the cell population they should be reduced by 0.1899 at least. I choose the value 0.2 for simplicity, thus the new probability of cell division set for hypoxia affected cells is shown in Table 5.5.5, and the growth curve this set of data produces is shown in Figure 5.5.3.2.

| Time | 0-24hr | 24hr-48hr | 48hr-72hr | 72hr -96hr | 96hr -120hr |
|---|---|---|---|---|---|
| Control condition probability of cell division for simulation | 0.42920 | 0.96362 | 0.70711 | 0.68990 | 0.61751 |
| Hypoxic condition probability of cell division for simulation | 0.22920 | 0.76362 | 0.50711 | 0.48990 | 0.41751 |

Table 5.5.5 Probability of cell division in my model for hypoxic group

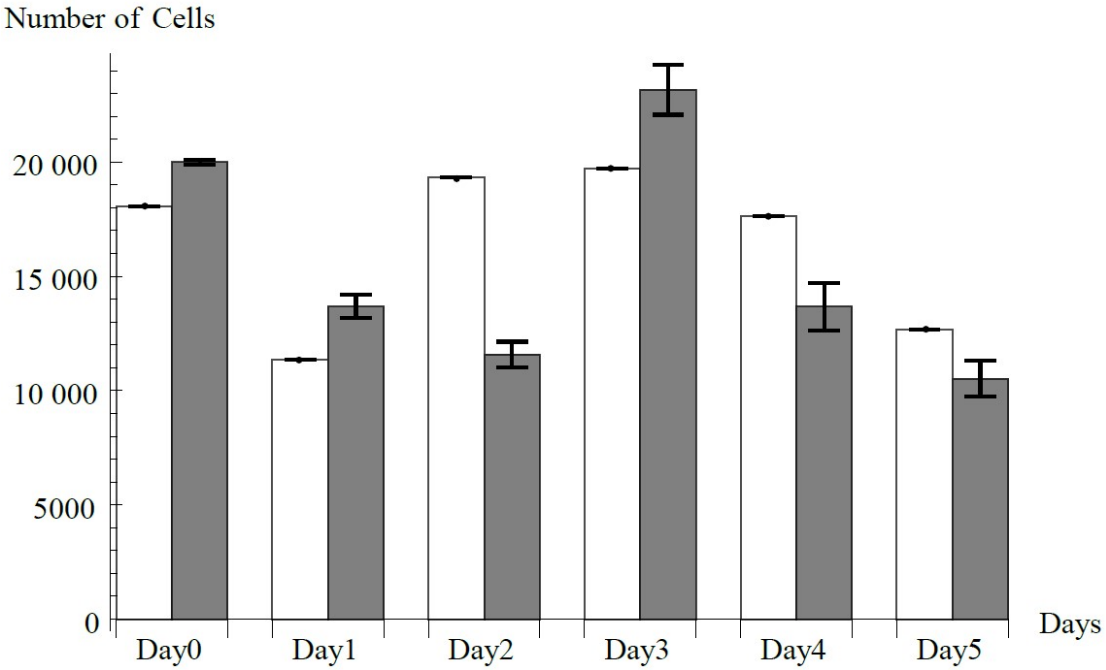


Figure 5.5.3.2 Hypoxia growth curve of my model and experiment result (with one type of agent). The white bars are simulation data and grey-scale bars are experimental data. As the simulation was run only once, there is not error bars for experimental data. The simulation and experimental data are very distanced on Day2.

Figure 5.5.3.2 shows that the growth curve from simulation and experiment are relatively close during 0-24 hr, then during 24-48 hr the simulation produces much more agents. Thus the simulation cell growth continues with an incorrect number of cells from 48-120 hr, so this it is not a good match either.

By observing the values in Table 5.5.5 I can predict the bad match of simulation result and experimental data, as the probability of cell division in the experiment in hypoxia group cannot be simulated by simply increasing or decreasing corresponding value in control group. In study in (Yijun Chen *et al*. 2009) shows a slowing of tumour cell proliferation with moderate hypoxia. Thus I decided to add a new type of agent to represent hypoxia affected cells with longer typical cell cycle time. I keep the old type of agent and let it maintain the original cycle length but further its probability of cell

division, while adding a new type which has a longer cycle but keep probability of cell division in previous chart, as shown in Table 5.5.6.

The new set of growth information is written into the setting file as a new type of agent, as Table 5.5.6 shows. The ‘cell type B’ is the new type of agent added to simulation (the original agent is called “cell type A”). As its cell cycle length is doubled, the value of ‘MaxAge’ and ‘AgeVariance’ is also doubled, as well as the end time that each value of probability of cell division takes effect. By adjusting the proportion of these two types in the starting population, I are able to delay the time where cell population begins to increase and produce growth curves shown in Figure 5.5.3.3.

| | Growth info for cell type A | | Growth info for cell type B | |
|---|---|---|---|---|
| MaxAge | 100 | | 200 | |
| AgeVariance | 10 | | 20 | |
| numberofphase | 5 | | 5 | |
| probability of cell division in phase 1 | Phase duration from hour 0 to hour 20 | 0.22920 | Phase duration from hour 0 to hour 40 | 0.42920 |
| probability of cell division in phase 2 | Phase duration from hour 20 to hour 40 | 0.76362 | Phase duration from hour 40 to hour 80 | 0.96362 |
| probability of cell division in phase 3 | Phase duration from hour 40 to hour 60 | 0.50711 | Phase duration from hour 80 to hour 120 | 0.70711 |
| probability of cell division in phase 4 | Phase duration from hour 60 to hour 80 | 0.48990 | Phase duration from hour 120 to hour 160 | 0.68990 |
| probability of cell division in phase 5 | Phase duration from hour 80 to hour 100 | 0.41751 | Phase duration from hour 160 to hour 200 | 0.61751 |

Table 5.5.6 Parameters relative to cell cycle and probability of cell division of two types of agents

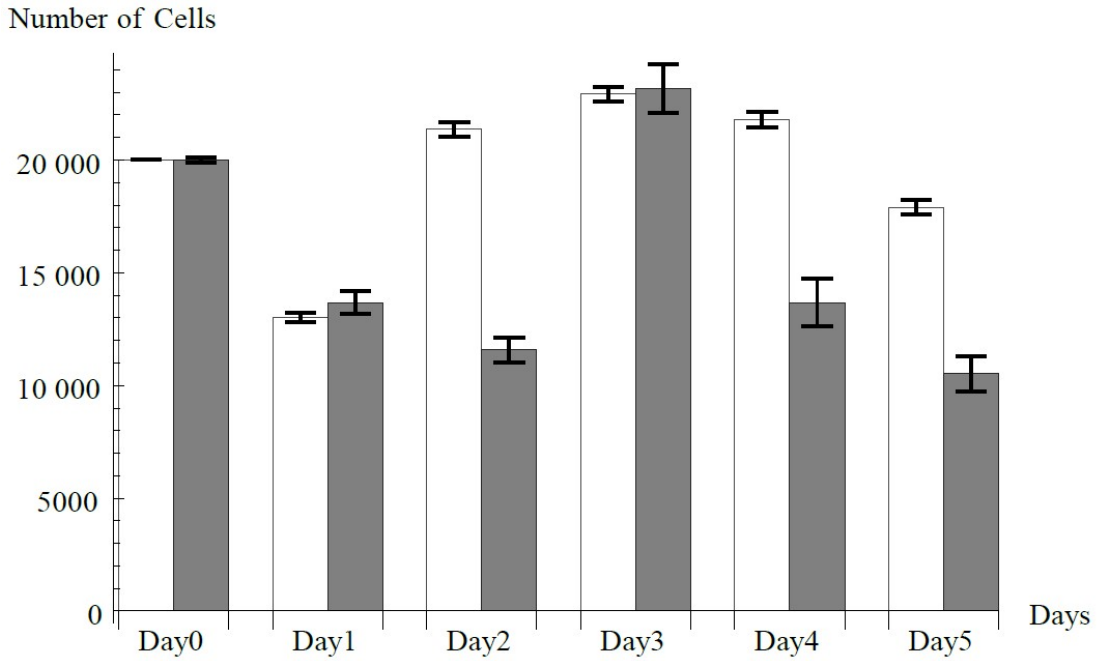



Figure 5.5.3.3 Hypoxia growth curve of my model and experiment result (with two types of agents) The value of simulation data is average of 10 runs, and the experimental data is from 3 counts. With two cell types, although they are still distanced on Day2, Day4 and Day5, the simulation and experimental data now match on Day1 and Day3.

Compared to Figure 5.5.3.2, the growth curve in Figure 5.5.3.3 is a better match. I can see from Figure 5.5.3.3 that although for most time points the simulated cell population level is close to the experimental results, it is still much larger than the experimental observation at the 48 hour point. It can be seen from the probability of cell division in Table 5.5.5 that for both cell types the probability of cell division is higher than 0.5, which causes the cell population to increase. However, the cell population decreases during this time in the experiment, which requires a probability of cell division lower than 0.5. As all the probability of cell division values must be altered equally under an assumption of constant drug effect, to solve the problem all the values should be decreased until the value for 24–48 hr is lower than 0.5, which will make the probability of cell division for 0–24 hr lower than zero and out of its possible range, and this means that this problem cannot be solved by tuning the

probability of cell division. A possible solution to this fitting problem is to tune the cycle length of the new cell type, until it remains in its first two cycles between 24 and 48 hr. As the old cell type is in 3rd-4th cycle during that time, the combination of probability of cell division of both types may fit the simulation growth curve to the experimental one. This requires further work to determine the proper cycle length and proportion of the two types in the starting condition. However, the results as they stand are a sufficiently good fit.

### 5.5.4 Predicting combined 5-FU and hypoxia population

To test independently the extensions to the model included to reflect 5-FU and hypoxia mechanisms of action, I modelled and compared the simulation output with experimental data for the condition combining 5-FU and hypoxia. This means that I preserve all parameters from the previous experiments: (1) two types of cells and their cell cycle and probability of cell division setting with which I simulate the effect of hypoxia; (2) the increase in chance of cell cycle arrest chance as used in the 5-FU affected simulation. Crucially, there is no new mechanism added and no data from experiment of combination of 5-FU and hypoxia is used.

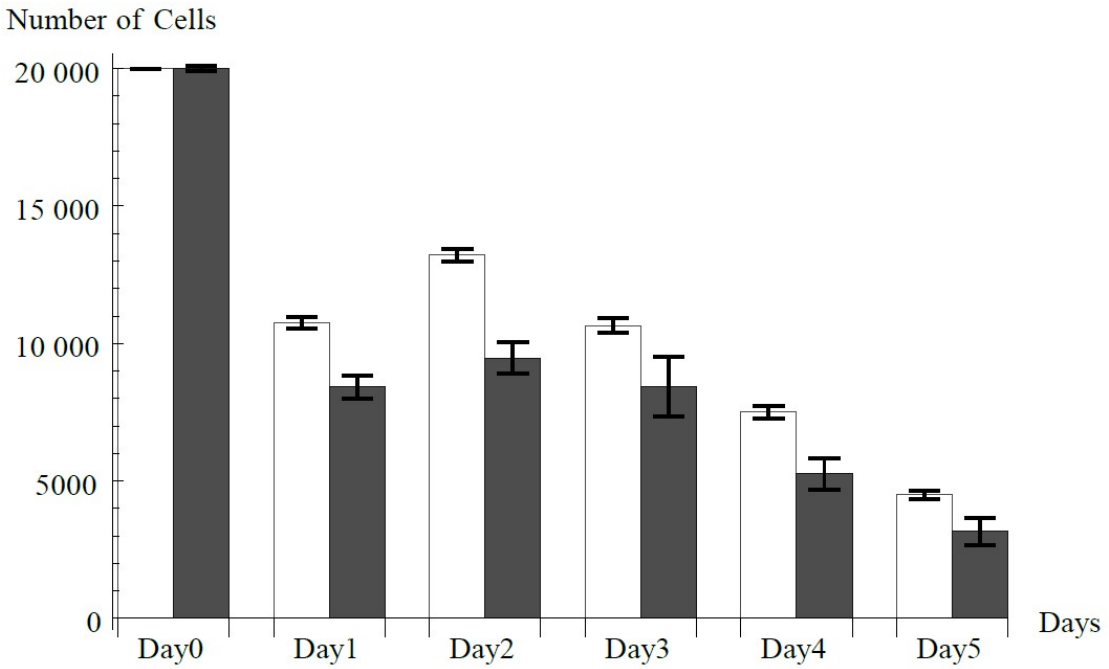



Figure 5.5.4.1 Simulations with combination of 5-FU and hypoxia.5% arrested agents compared with experimental data: the simulation bars are in white and the experiment bars are in grey. The value of simulation data is average of 10 runs, and the experimental data is from 3 counts. The simulation and experimental data do not match but can be considered as a good result.

Figure 5.5.4.1 shows that, although the simulated and experimental data do not match, they are close enough to be considered as a good result, where with no other model adjustment I are able to predict periods of population growth and decline. I note some differences in the rates of decline and increase in the earlier stages of the simulation. According to previous simulation results, this difference is most likely driven by the hypoxia effect, because the 5-FU simulation tends to have similar dynamics over all time points, while the hypoxia simulation has higher variability during the period of most difference in Figure 5.5.3.3. I can see that the combined effects show a reasonably good fit to the experimental results, and I have reason to believe that, by improving the fit of the hypoxia simulation, I are able to improve the combined simulation at the same time.

# Conclusion and discussion

In this chapter, the model developed in Chapter 3 has been revised to provide a framework for modelling cancer cells. The main amendment is the inclusion of a cell cycle and probability of cell division, which are necessary since they are important factors of cell group population which needs to be calibrated before the group pattern formed by cancer cell.

The cell cycle representation is phenomenological, and in this chapter it is calibrated to the population growth curve in the control condition experiment. According to the effect of 5-FU on cell cycle (Choudhary, *et al*, 2012), by adding in an extra death rate, the mechanism of 5-FU is added to the initial model. With 5% extra death rate, the simulation of model shows a close result in comparison with experimental data.

However the same method cannot be used to add the mechanism of hypoxia, as it not only increases the death rate of cells, but also slows their growth and hence extends the length of typical cell cycle (Yijun Chen *et al* 2009, Ortmann *et al* 2014). The mechanism of hypoxia is modelled by including two cell types: one with higher death rate and normal cell cycle, one with normal death rate and longer cell cycle. Comparing with simulation with single-cell-type, simulation of mixed cell type produces better result.

After calibration of model separately with both 5FU and hypoxia experiments, the settings of these two parts are combined together to predict the cell growth with combination experiment. The simulation of combination model shows that, although it does not match the experimental data, it is able to show a similar trend of change of

cell population. From this I consider the model can produce proper cell population for spatial pattern study which will be discussed in Chapter 6.

In the next chapter I repeat this approach of calibrating to the cell spatial patterning with the built-in cell cycle calibrated in this chapter: the control condition is calibrated initially, followed by introducing plausible mechanisms to reflect 5FU and hypoxia separately; at last the combined treatment of hypoxia and 5FU is predicted.

# Chapter 6 Predicting cell distribution curve

## Introduction

In chapter 5, I mapped the population growth curve of experimental data to the model, which ensures that the simulation containing correct number of agents. In this way the inter-cell model is ready to be used to study the features of cell distribution.

The aims of this chapter are to:

(1) Find a method to characterise spatial experimental data;

(2) Using the method found in (1), analyse the data gathered from growth experiments to find the key properties of the distribution of cells under various environment conditions;

(3) Find a method to link the spatial distribution of cells to the parameters of the physical model where the model produces similar distributions;

(4) Compare the parameter setting of physical model corresponding to each experiment, finding the differences between experiments and explain those differences in a biological context.

## 6.1 Time-lapse image process

First of all, the data gathered from the experiment is in form of time-lapse images. According to Section 5.3, the images are taken with an interval of 10 minutes from when the HCT cells are first established in the petri dish. All of images of each growth experiment are focused on a fixed area of part of one of the 96 wells. In this section I review the process to extract the information relating to cell distribution from the time-lapse images, which is Aim (1) in the introduction of Chapter 6.

To characterise the cell distribution is to characterise the distribution of cell positions. Thus to find the key properties of cell distribution, the first step is to register the positions of cells in the time-lapse images. The set of positions is then described by a point process, a spatial statistic method that can measure a distribution in space and provide quantitative descriptors of that distribution.

### 6.1.1 Pre-process

In the growth experiments, the experimental setup was not ideal due to equipment limitations. The number of cells was counted regularly, a process that perturbed the experimental setup and so may slightly affected the relative position of petri dish and camera. Thus some of the time-lapse images are taken under slightly different conditions. Some images vary in brightness, and some blur because of lose of focus. Also on each image there is a large bright area and two smaller dark areas which are obviously noise. For individual cells to be recognised by image processing software, these problems need to be resolved.

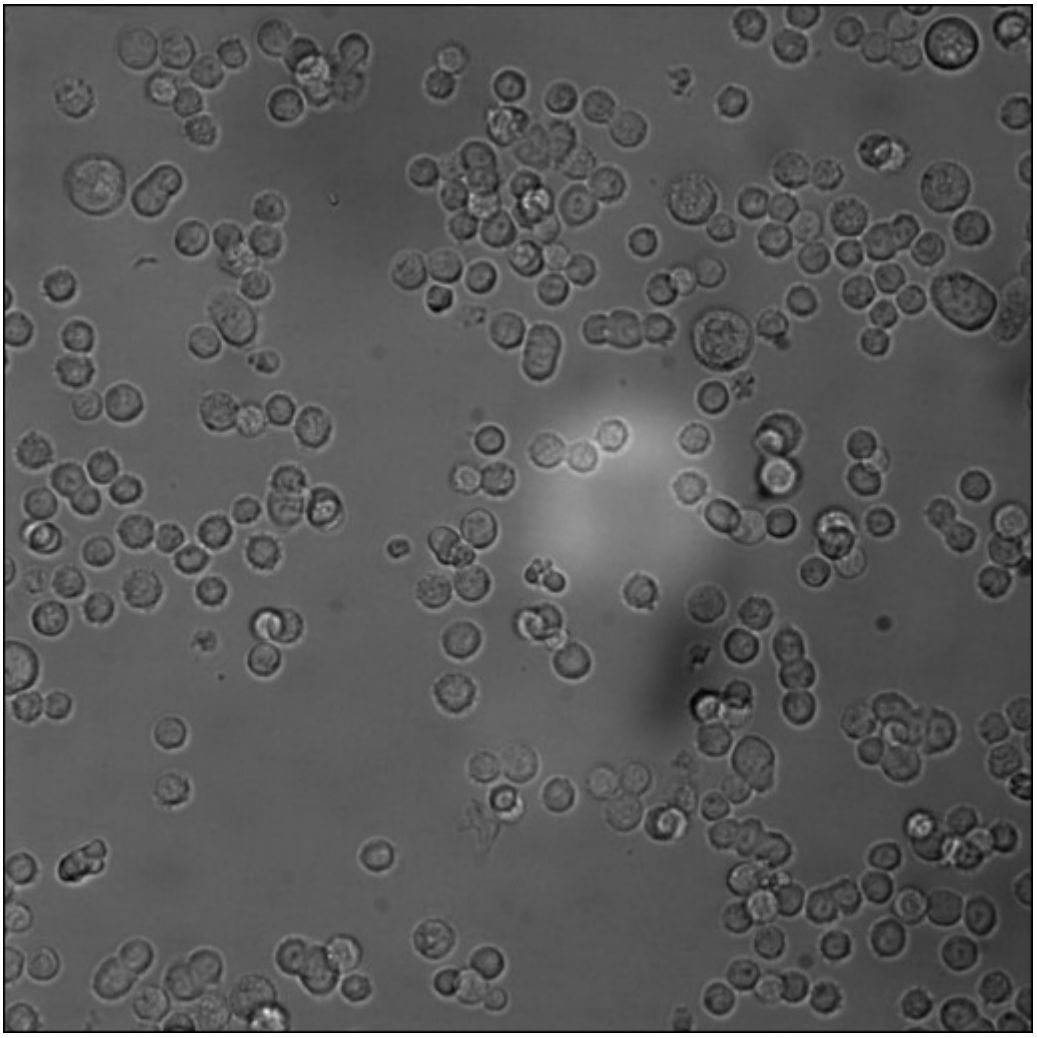
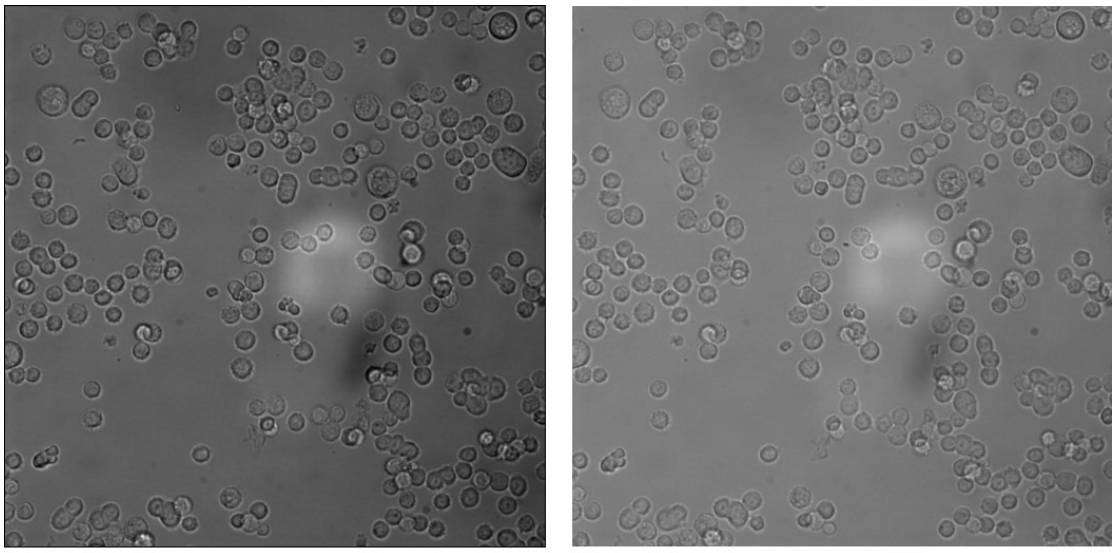


Figure 6.1.1.1 Blurred image with bright spot in the middle and dark spots on bottom-right and top-right part. The aim of image-process is to filter the bright spot and dark spots.

There are three operators in the software ImageMagick (http://www.imagemagick.org/) that can be used to help solve these problems. One of them is the "-level" operator. This operator requires two input values for thresholding upper and lower boundaries: 'black_point' and 'white_point'. Any part that is darker than the 'black_point' is set to be black; and any part that is brighter than the 'white_point' is set to be white. After that "-level" operator scales the greyscale of all the rest part linearly (http://www.imagemagick.org/Usage/color_mods/#level). In this way I are able to increase the contrast of image, as Figure 6.1.1.2 shows. Also, for a group of images taken from one experiment, by assigning the same 'black_point' and 'white_point',

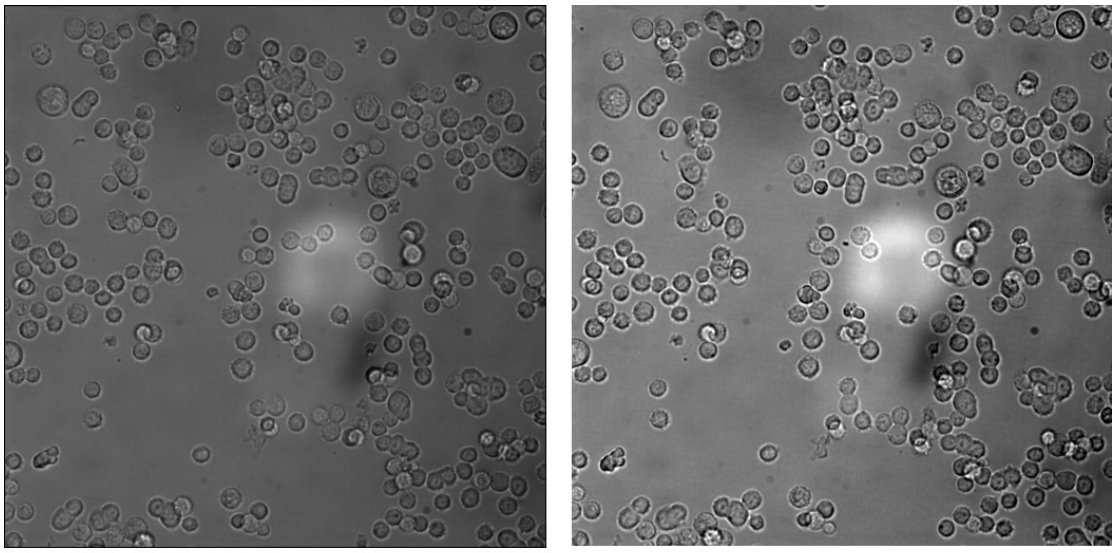
the whole set of images can have an even brightness.



(a) (b)

Figure 6.1.1.2 Before (a) and after (b) increase of contrast (-level).

Another operator is "-normalize", which expands the distribution of greyscale of each image to 0~255 (http://www.imagemagick.org/Usage/color_mods/#normalize). The darkest part of each image is converted to greyscale 255 and brightest part of each image is converted to greyscale 0, and for the rest part the greyscale is scaled so that any value between 0 and 255 can be found. In this way, the dark part of the original image looks darker, and bright part looks brighter, as Figure 6.1.1.3 shows.



(a) (b)

Figure 6.1.1.3 Before (a) and after (b) normalisation (-normalize)

Theoretically the detail of the darkest 2% and brightest 1% are lost with "-normalize" operator, which means the darkest 2% of greys in original images are turned to black and brightest 1% of greys are turned to white. From Figure 6.1.1.3 I can see that this loss of detail does not significantly affect the quality of image, and importantly the contrast of the image is enhanced.

The last operator is “LoG” (http://www.imagemagick.org/Usage/convolve/#log). It is a high pass filter to get rid of the bright spot and dark spots on the image. The image before and after operation are shown in Figure 6.1.1.4.

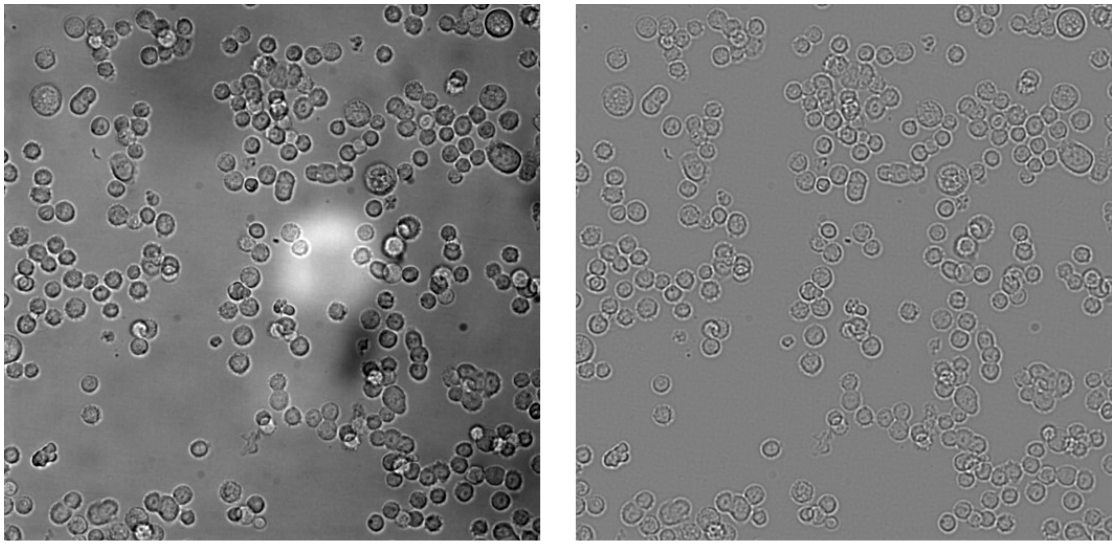



Figure 6.1.1.4 Before (a) and after (b) high pass filter (-LoG)

An important consideration is the order of application of these operators. Because the "-normalize" operator increases the contrast within an image, the three largest noisy areas (the dark and bright regions) will be more pronounced and thus harder to remove if "-normalize" is used before “-level”. On the other hand, the operations that increase the contrast of image, such as "-normalize" and “-level”, can produce noise

during rescaling of the greyscale, as Figure 6.1.1.5 shows. Therefore the “-level” operator is the first to be applied, followed by "-normalize; “LoG” is the last operation.

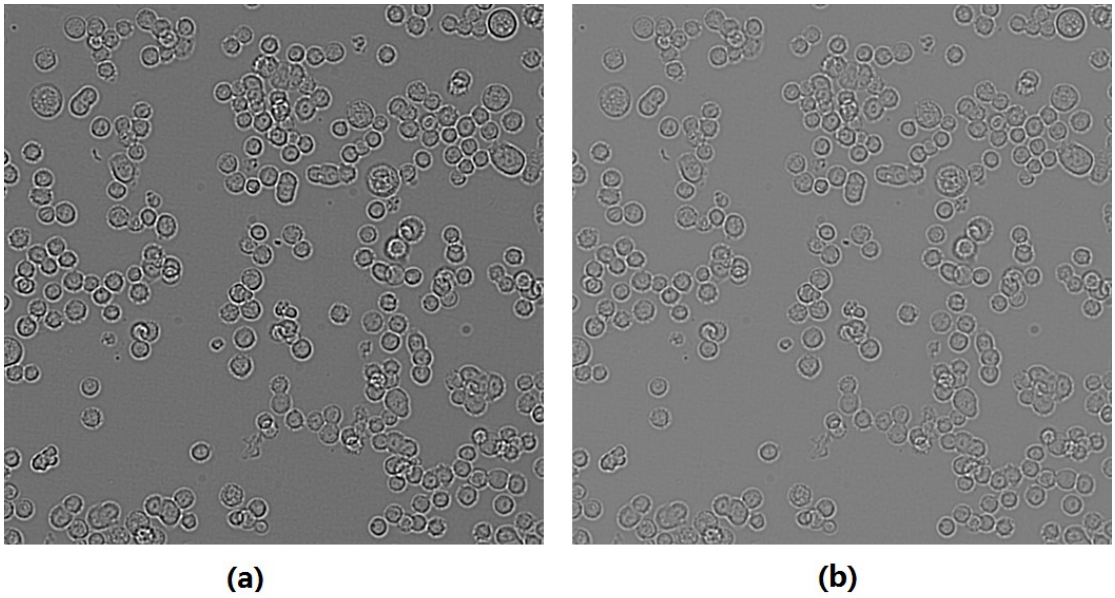



Figure 6.1.1.5 Result of high filter pass then normalise (a) and normalise then high filter pass (b). Cells are more obvious in (a).

## 6.1.2 Recognising cell outlines in CellProfiler

CellProfiler is a free to use cell image analysis software, available from [http://www.cellprofiler.org/]. It provides functionalities to allow recognition of the edge of cells and recording of cell positions through a graphical user interface.

Each processing function is packaged as one module with a separate setting up process. Several modules can be combined in a specified order, so that the output of a previous module can be used as an input of later module. In this way, several modules operating in a fixed order form a pipeline that can be saved or loaded in CellProfiler. CellProfiler automatically runs the all modules in a pipeline with that fixed order.

One of the modules to load an image is always set as the first module in the pipeline. In the ‘Image’ module, image files or folders are dragged to the interface. The second module in the pipeline is ‘Metadata’. The purpose of this module is to extract the pattern in the naming of loaded images, so that a similar name can be assigned to the corresponding output file. Then in module ‘NamesandTypes’, the image type is to be set as ‘Greyscale image’; also the group of images is assigned a name to be used in the following modules.

The module used after input modules is ‘IdentifyPrimaryObjects’. In this module, the edge of cells is identified from input images. An identification result is shown in Figure 6.1.2.1: the lines in green are accepted edges; the ones in red are unacceptable edges according to their size; and the yellow lines are unaccepted edge because they touch the edge of image.

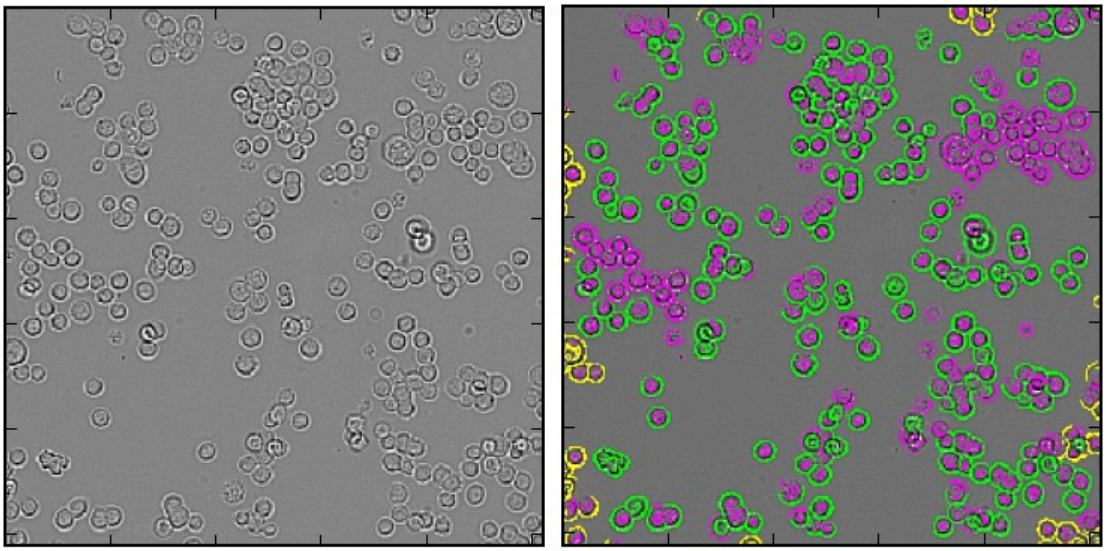



Figure 6.1.2.1 A time-lapse from experiment (a) being identified (b). The green parts in (b) are accepted edges of cells; the red and yellow parts in (b) are unacceptable edges of cells. Note the red part on up-right corner, where a cluster of cells is failed to be identified. This problem is discussed in Section 6.1.3.

It is recognised that for some images, different values of parameter are needed in the 'IdentifyPrimaryObjects' module, which are set manually. For detail of actual value, please refer to the Appendix D. This non-systematic threshold adjustment is not an ideal image recognition process. However, image recognition is not the aim of this thesis: it is a pro-analysis step to retrieve data from time-lapse images for further study, thus the driving concern here is that data generation. A more controlled experimental system and more time spent investigating image analysis algorithms would overcome this limitation in the cell recognition process.

After module 'IdentifyPrimaryObjects', module 'ExportToSpreadsheet' is called to save position of identified cells to text files. I need to set the output data is from module 'IdentifyPrimaryObjects', the format of name of output file which is from 'Metadata' module, and how each column of data is separated.

### 6.1.3 CellProfiler parameter calibration

In Section 6.1.2 it is claimed that the parameters are chosen in order to produce the best result. Firstly the parameters are tested and values of parameters and number of cells found are recorded. The values are chosen so that the most number of cells are recognised while the lease sundries are recognised as cells. To calibrate the parameters in CellProfiler, the identified cell positions are drawn back to the original image and compare how good they fit. Before drawing cell positions onto time-lapse images, the cell position data needs to be filtered. Because sometimes the output file from CellProfiler contains duplicated cells or two cells with very close positions. In

addition the output from CellProfiler also contains image ID and cell ID that need to be skipped. A separate c++ program DrawCellPosition.sln has been implemented to remove the unnecessary columns from output data and filter out cells that are too close or are duplicated. After that the filtered positional data is drawn as a white cross on a copy of time-lapse image to show filter result.

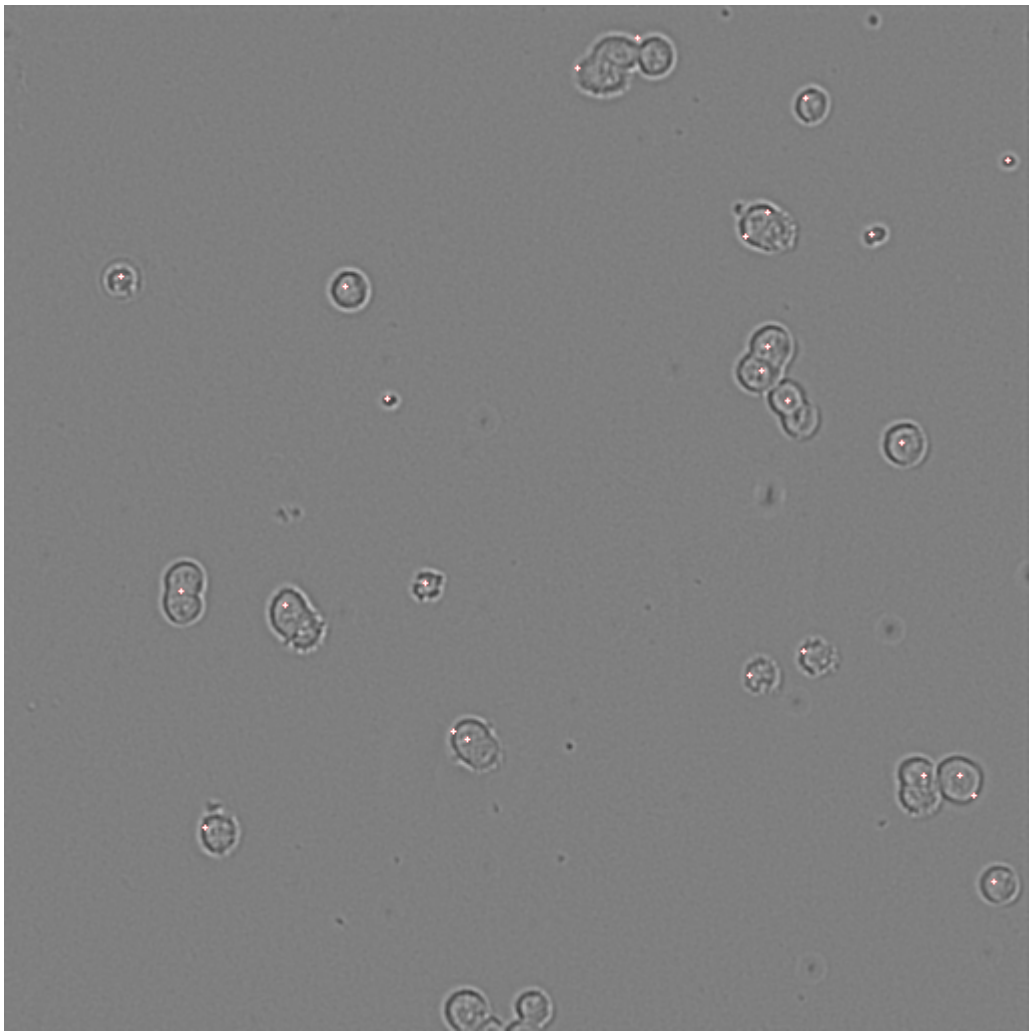


Figure 6.1.3.1 time-lapse image with identified cell position drawn as white cross

By plotting the recognized cell position on the original image, it is found that cluster of cells cannot be properly recognised. In most common circumstances, the outline of clustered cell group can be recognized, while the edge of each cell within cluster cannot be distinguished and thus the cluster is considered as one single cell positioned

at the centre of the cluster.

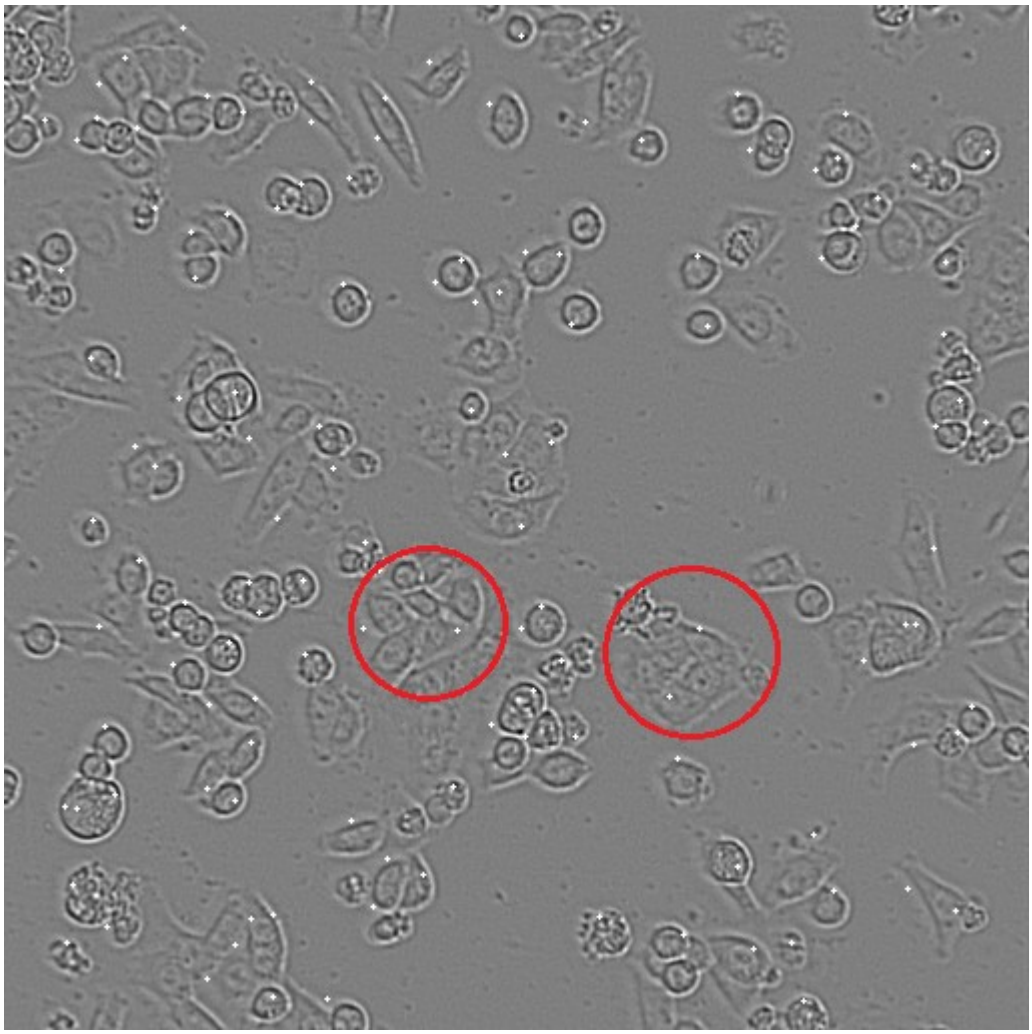


Figure 6.1.3.2 Cluster considered as single cell

To investigate how much error is introduced to spatial distribution measurement by failure to recognise cells in clusters, a test is carried out in Section 6.2.3 to compare the difference of spatial distribution between the conditions that cells in clusters are recognised and are not recognised.

### 6.1.4 Cell recognition problems and manually processing

In practice, by using the approach outlined in Section 6.1.3, it was found that

CellProfiler cannot distinguish among individual cells within a cluster. Although the edge of cells in clusters is clear for human eyes to identify, the edge is not a closed circle due to lighting condition and in CellProfiler non-closed circle cannot be identified as a cell. This lack of differentiation means that both the total number of cells and the measurement of the distance between cells are imprecise. In order to determine how much error recognition problems bring to the measured distribution, the position of cells of the same images (that has been processed by CellProfiler) are manually marked and recorded. The total number of cells identified by CellProfiler and by hand is shown in Table 6.2.3.1.

| | First image | Middle image | Last image |
|---|---|---|---|
| Number of cell in control group indentified by CellProfiler | 29 | 26 | 38 |
| Number of cell in control group indentified manually | 24 | 28 | 21 |
| Number of cell in 5-FU group indentified by CellProfiler | 89 | 85 | 111 |
| Number of cell in 5-FU group indentified manually | 243 | 240 | 192 |
| Number of cell in hypoxia group indentified by CellProfiler | 89 | 173 | 170 |
| Number of cell in hypoxia group indentified manually | 146 | 148 | 122 |
| Number of cell in combination group indentified by CellProfiler | 100 | 108 | 147 |
| Number of cell in combination group indentified manually | 140 | 144 | 143 |

Table 6.1.4.1 Number of cells identified by CellProfiler and by hand

From Table 6.1.4.1 I can see that the numbers of identified cells are similar for the control group; for other groups, the number of cells identified by CellProfiler is lower than manually identified data, especial for the 5-FU group. However, comparing with manually process, the CellProfiler can identify similar number of cells with manually for the last image of combination group, which means the performance of CellProfiler

is unstable.

To test how different identified cell numbers affects the measurement of cell distribution, the pairwise correlation function for both CellProfiler and manually recognized cell position is separately calculated and plotted, as Figure 6.2.3.1 shows.

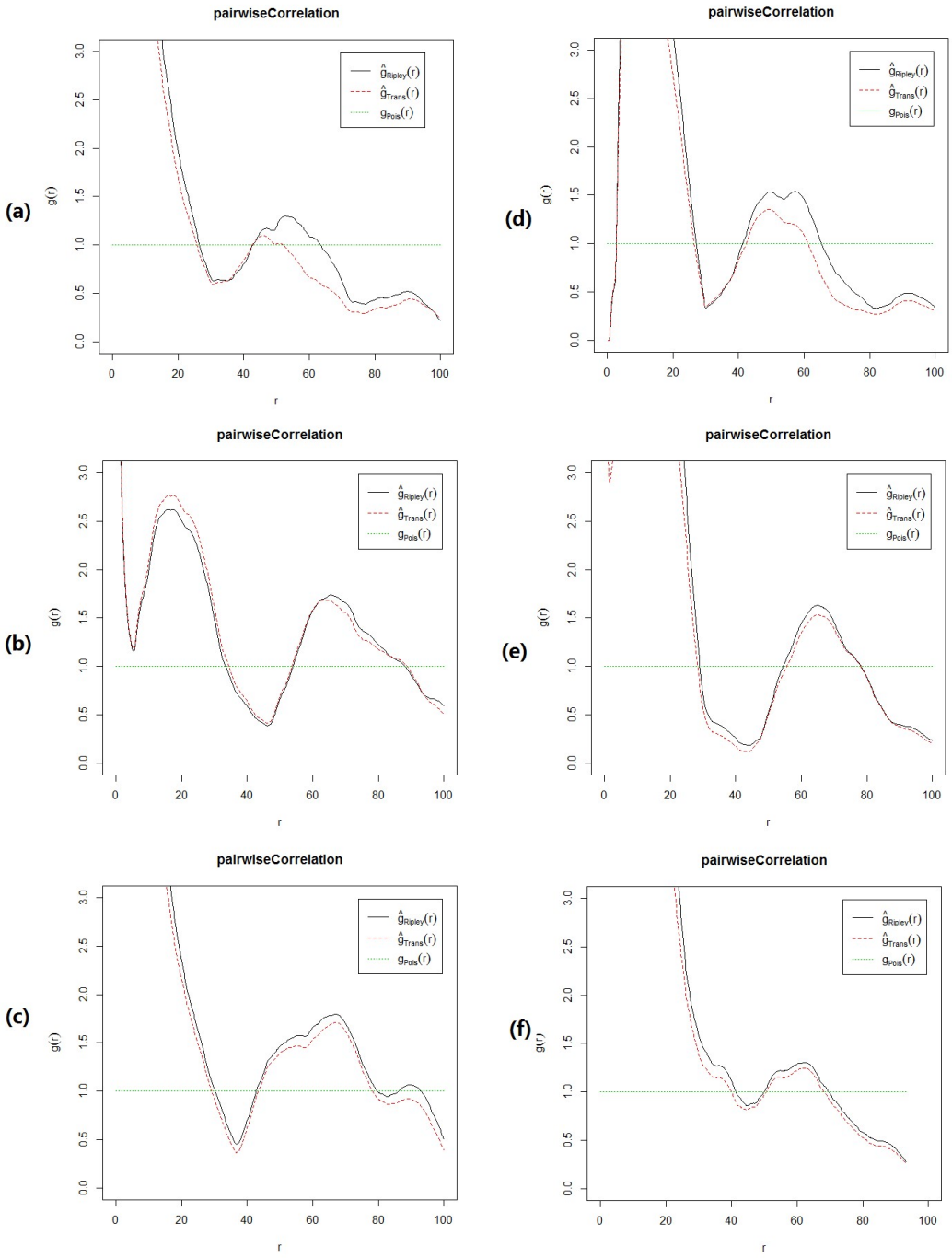



Figure 6.1.4.1 Pairwise correlation of the HCT wild type control group. (a) - (c) are

generated with first, middle and last image recognised by CellProfiler, and (d) - (f) are generated with first, middle and last image manually recognised.

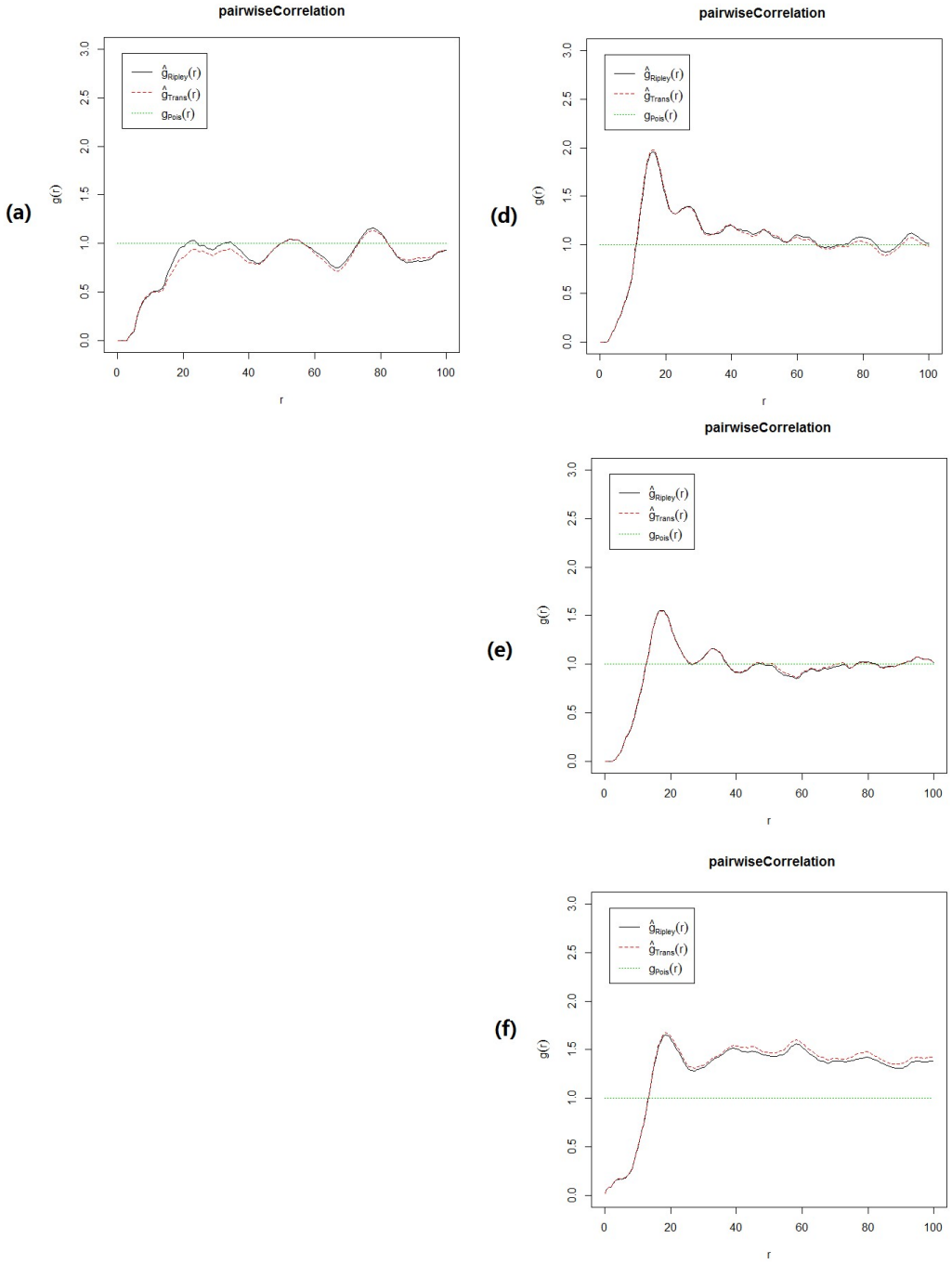



Figure Figure 6.1.4.2 Pairwise correlation of HCT wild type with 5-FU. (a) is generated with the first image recognised by CellProfiler, while (d) - (f) are generated with the first, middle and last image manually recognised. Note that (b) and (c) are missing. Because from the middle and end image of 5-FU group, the CellProfiler cannot identify enough cell to produce pairwise correlation curve.

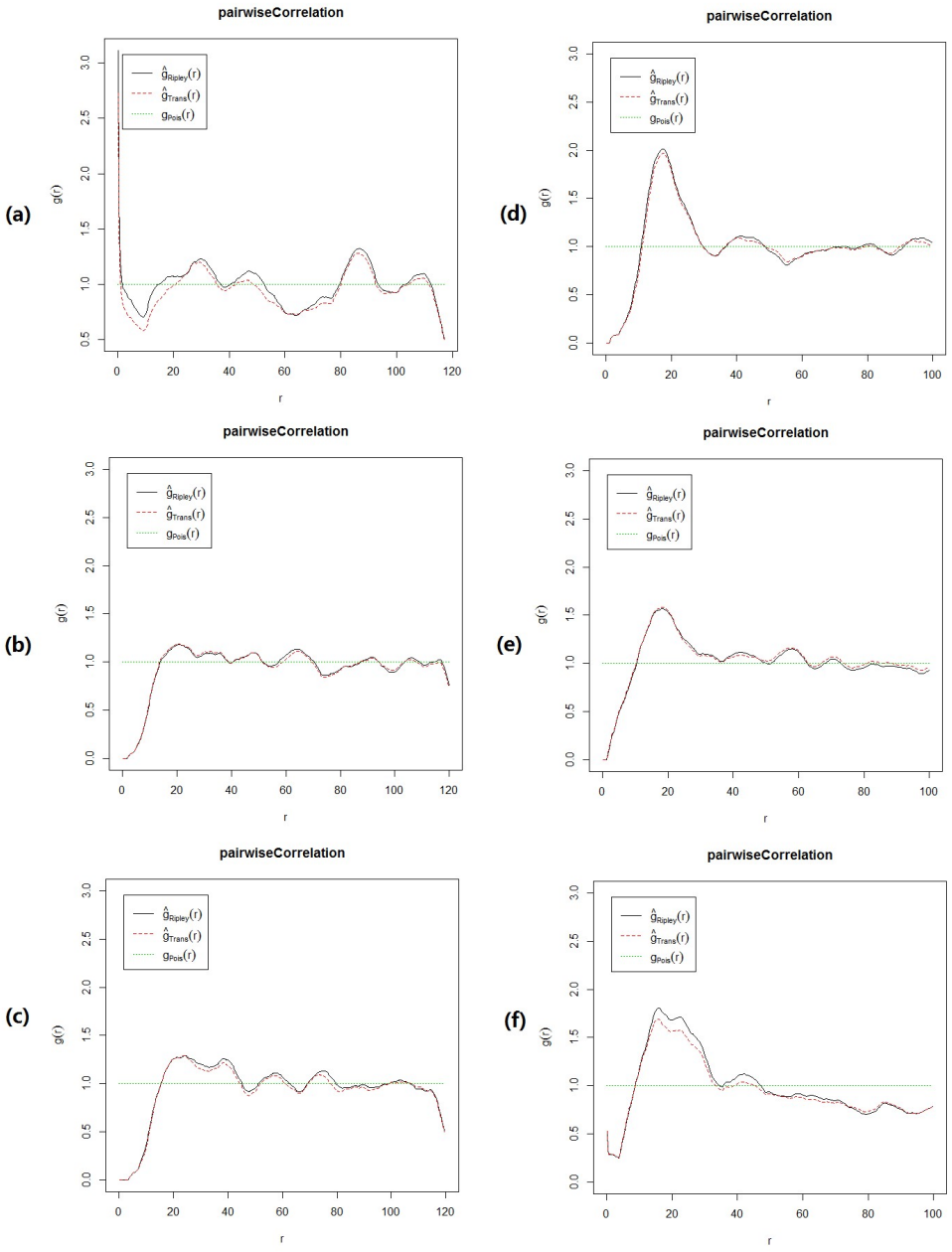



Figure 6.1.4.3 Pairwise correlation of the HCT wild type with hypoxia. (a) - (c) are generated with first, middle and last image recognised by CellProfiler, and (d) - (f) are generated with first, middle and last image manually recognised.

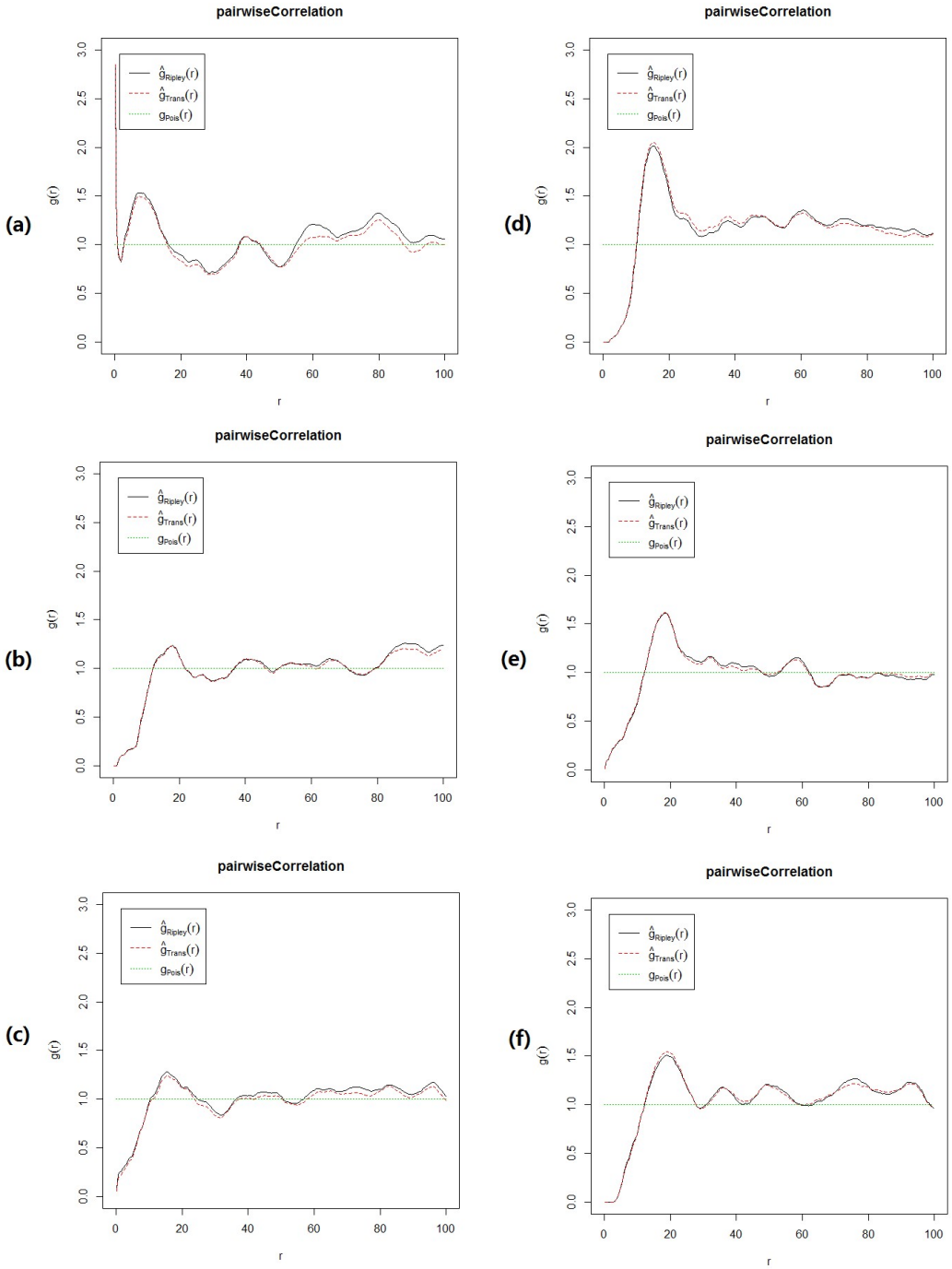



Figure 6.1.4.4 Pairwise correlation of the HCT wild type with combination of 5FU and hypoxia. (a) - (c) are generated with first, middle and last image recognised by CellProfiler, and (d) - (f) are generated with first, middle and last image manually recognised.

Comparing the pairwise correlation function in Figure 6.1.4.1, Figure 6.1.4.2, Figure 6.1.4.3 and Figure 6.1.4.4, it suggests that for the control group, there is not much difference between CellProfiler generated and manually generated data. However for

all the other experimental groups, the CellProfiler processed data loses the first peak of pairwise correlation, which is caused by the loss of data of the close distance between cells that are in the same cluster; and the overall height of the CellProfiler curve is slightly lower than the manually processed data, which is caused by the reduced total number of recognized cells. The reason that the control group is not significantly affected is that in image of control group there are much fewer clusters, and within each cluster the number of cells is much lower.

Finally, due to the imprecise result produced by CellProfiler, I choose to use manually process data for later analysis in all the following sections. Also the conclusion of this chapter is obtained from the manually process data. While not an ideal solution, image analysis is not a major focus of this thesis and so a manual analysis was judged an acceptable solution (it is discussed in Chapter 7, Discussion).

## 6.2 Fitting experimental parameters to the starting point of the simulation

Before being analysed, the experimental data and simulation result should be measured at same scale. The diameter of cells in the control experiment is observed to be 10 $\mu m$; note in the time-lapse image the diameter of one cell is about 10 pixels. Thus I consider in time-lapse image, each pixel represents a $1\mu m \times 1\mu m$ area, and on spatial statistic analysis curves each unit represents length of $1\mu m$. As the sizes of the images are $512 \times 512$, each image represents a $512\mu m \times 512\mu m$ area. Similarly the result of a simulation needs to be of same scale. The output position of agents is then expressed in terms of unit length. To unify the experimental and simulation scales, the

output value should multiple value of unit length, then divide by $1\mu m$.

The result of simulation is measured by unit length, which equals $7.5\times10^{-5}$ m. To convert the simulation result so that it is measured in micrometers, which equals $10^{-6}$ m, the position data needs to be multiplied by $7.5\times10^{-5}$ then divided by $10^{-6}$, as shown in Equation (165).

$$NewPosition = \frac{OldPosition \times unitlength}{micrometer} \quad (1)$$

For each experimental group, I select images that depict the beginning of the experiment, mid-point and end of the experiment to carry out image recognition of the edges of cells, which result is later analyzed and compared with point process.

The number density, on the other hand, should be of same value in the start point of both simulation and experimental data regardless of scale, so that the simulation starts from the same condition of experiment.

As the intensity of experimental data is known at the starting point of simulation, the number of agents should be the same with number of cells in the first image from each experiment. In the simulation, there is a setting for the density of agents, which translates to the number of agents per unit area. The number density of agents is calculated as follows:

number density of agents = number of cells in experiment / area of substrate in simulation (2)

New unit length: 7.5e-005 m

New unit time: 60 seconds

Area of simulation substrate: 6.83 unit length * 6.83 unit length

So: 6.83*7.5e-005 m = 512.25 e-006 m -> experiment area $512\mu m \times 512\mu m$

| | number of cells | Corresponding number density of agent |
|---|---|---|
| 5-FU group | 243 | 5.21 |
| Hypoxia group | 146 | 3.13 |
| Combination of 5-FU and hypoxia group | 140 | 3.00 |

Table 6.2.1 number density of agents in three experimental groups

The values of parameters such as unit length etc. are calculated for cancer cell growth as well; see Section 3.9 and Appendix B for full detail.

| Name of parameter | Value |
|---|---|
| Unit length $\Delta x$ | $7.5\times10^{-5}$ m |
| Unit time $\Delta T$ | 1 minute |
| Density of cell | $2000 kg/m^3$ |
| The proportion of typical cell size and unit length | Three semi-radii of ellipsoidal agent: $1.25\times10^{-6}$ m, $5.0\times10^{-6}$ m, $5.0\times10^{-6}$ m |
| Dynamic viscosity of fluid in which cells are moving | $1\times10^{-3} Pa\cdot s$ |
| Constant for contact force $K$ | $25.6\times10^{12}$ |
| Constant for adhesion force $\varepsilon$ | $0.6\times10^{6}$ |

Table 6.2.2 Value of parameters used in simulation of cancer cell growth. . Note that the values of $K$ and $\varepsilon$ are their dimensional value $\tilde{K}$ and $\tilde{\varepsilon}$. As $\tilde{K}$ and $\tilde{\varepsilon}$ do not appear in Chapter 3, I use $K$ and $\varepsilon$ here for a better understanding

In the simulation, the initial position of all the agents is randomly generated. To test the functionality of randomised position generation, the K-function curve of starting condition of simulation is plotted and compared with K function curves of the corresponding experiment, as shown in Figure 6.2.1, 6.2.2, and 6.2.3.

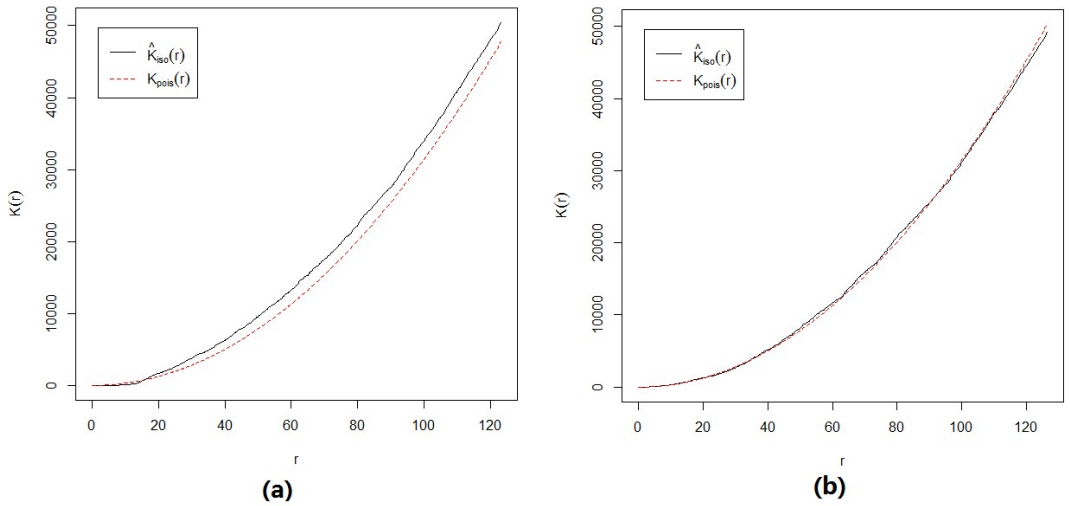



Figure 6.2.1 (a) K function of 5FU experiment group, (b) K function of 5FU simulation

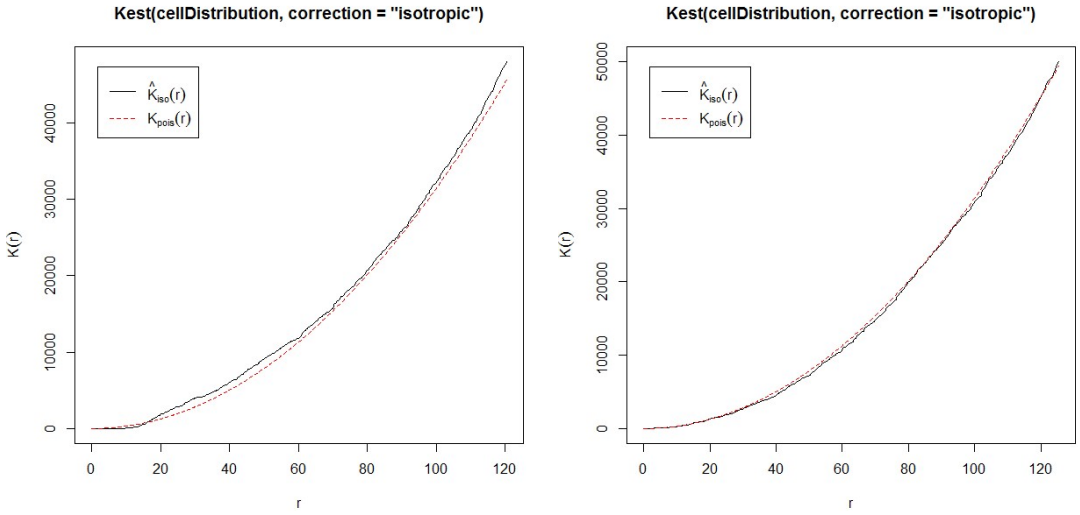



Figure 6.2.2 (a) K function of hypoxia experiment group, (b) K function of hypoxia simulation

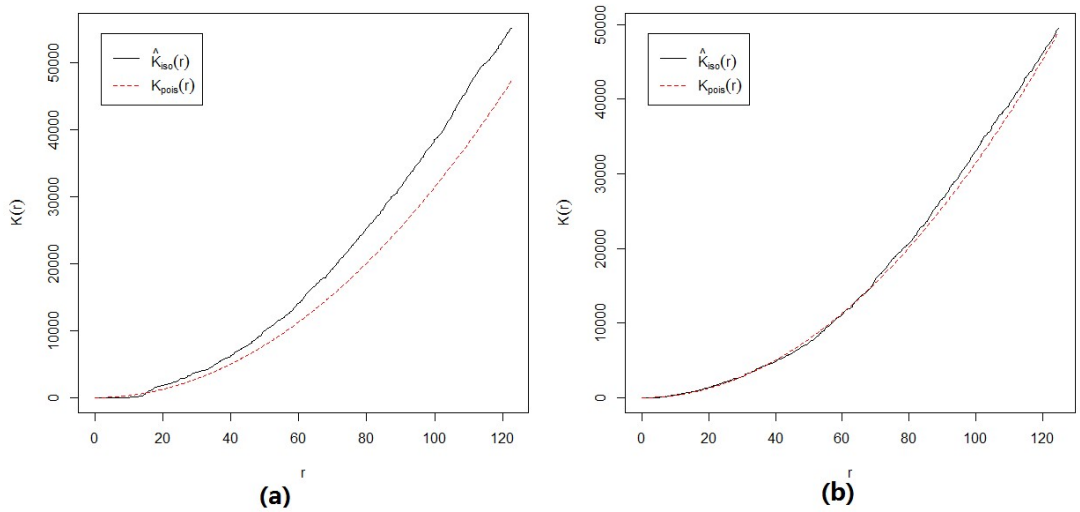



Figure 6.2.3 (a) K function of combination of 5FU and hypoxia experiment group, (b) K function of combination of 5FU and hypoxia simulation

I also fit the starting point of 3 simulations to stationary Poisson process. The output number density of simulation is shown in Table 6.3.2.

| | number density |
|---|---|
| Experiment group: 5-FU | 0.00099 |
| Simulation: 5-FU | 0.00094 |
| Experiment group: Hypoxia | 0.00061 |
| Simulation: Hypoxia | 0.00057 |
| Experiment group: combination of 5-FU and hypoxia | 0.00058 |
| Simulation: combination of 5-FU and hypoxia | 0.00055 |

Table 6.2.3 number density of 3 experiments and start point of simulation

From the K-function curve and number density value, it is demonstrated that, at starting point, the simulation is able to generate similar distribution with experiment. The scaled simulation is ready to begin with parameter fitting.

## 6.2.1 Pairwise correlation and 4 main parameters that affect the distribution curve

Before full calibration of my model, and to inform the parameter fitting process, I

carried out instructive simulations to explore the relationship between model parameters and agent distribution.

By observing the shape of the pairwise correlation curve, I are able to make initial assumptions about the link between the form of the pairwise correlation curve and the value of the parameters to the inter-cell interaction model.

For example, Figure 6.2.1.1 shows the starting point of an explorative simulation, in which the pink particles are agents and are randomly generated, the blue part is the back group representing the substrate, and the distribution of agents is shown with a top-down point of view.

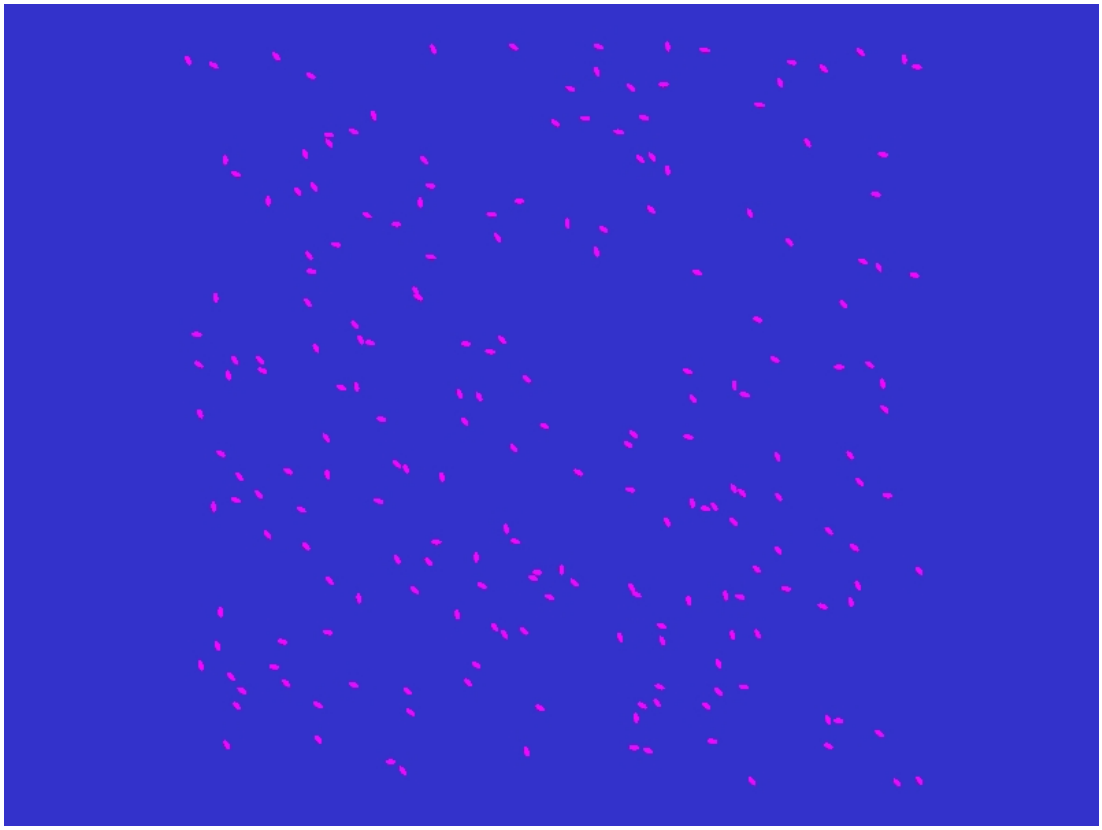


Figure 6.2.1.1 Starting point of simulation. Pink dots are agents which position and directly are randomly generated.

In this section I will explore a couple of key parameters in the model by running simulations with various parameter settings and compare the simulation result. Then it is known how each parameter affect the pattern formed by agents. The positions of agents of each simulation result are used to generate pairwise correlation curve. Then the effect of each parameter is linked with change of shape of curve (hopefully).

With definition of pairwise correlation function discussed in Section 6.2.1, I give meanings of each part of pairwise correlation curve as Figure 6.2.1.2. Note that the x-axis is unit distance between pair of agents, and y-axis is value of pairwise correlation function at each distance. The green dotted line is where pairwise correlation function $g = 1$. The part of curve above dotted line means the agents are clustered at this distance; the part of curve beneath means the agents are separated at this distance; the part of curve on dotted line means the agents are randomly distributed at this distance. The distribution range means the longest range on which another agent is found. The highest point is a value generated during process of calculation of pairwise correlation function and I believe it represents the degree of gathering of agents. These parts are considered as key feature of each curve and will be compared in the following discussion.

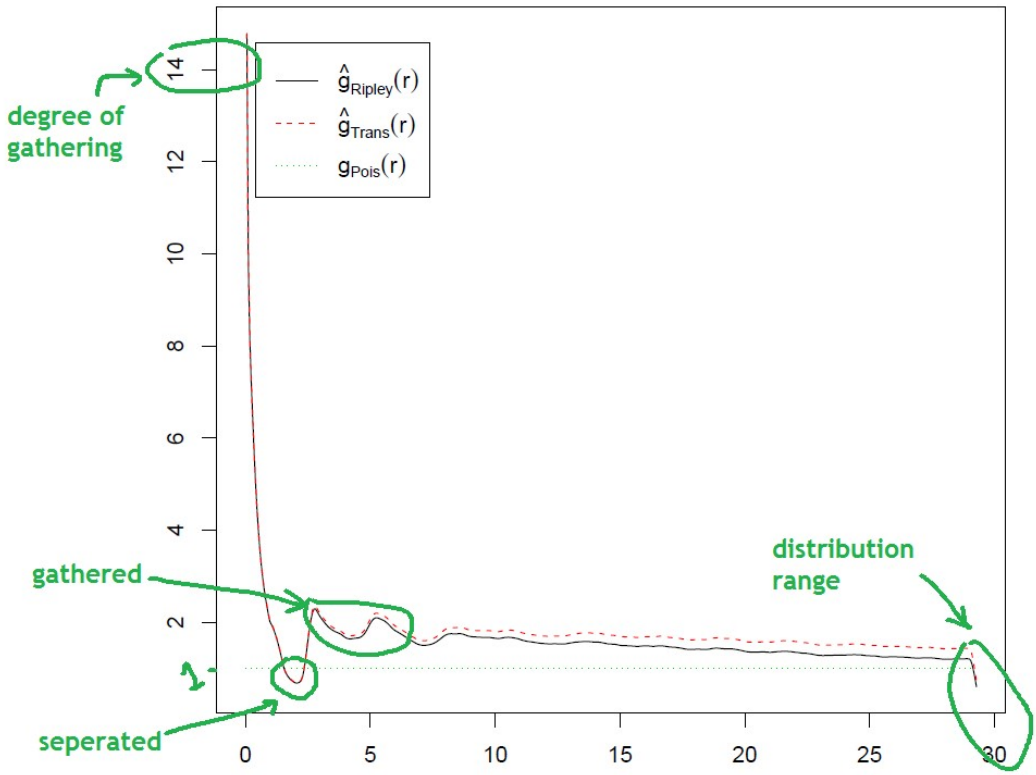



Figure 6.2.1.2 Interpretation of parts of pairwise correlation function.

## The range that adhesion (attraction) force takes effect

Firstly I consider the range that adhesion force takes effect. Note that the adhesion force attracts agents together, while the contact force repulse agents to prevent them from totally overlap. As the contact force affects agents only when they are in contact, the range of contact force is fixed and does not need to be compared. 4 simulations are run, with range of adhesion force set to 100, 10, 3 and 1 separately. The results are shown in Figure 6.2.1.3, 6.2.1.4, 6.2.1.5 and 6.2.1.6.

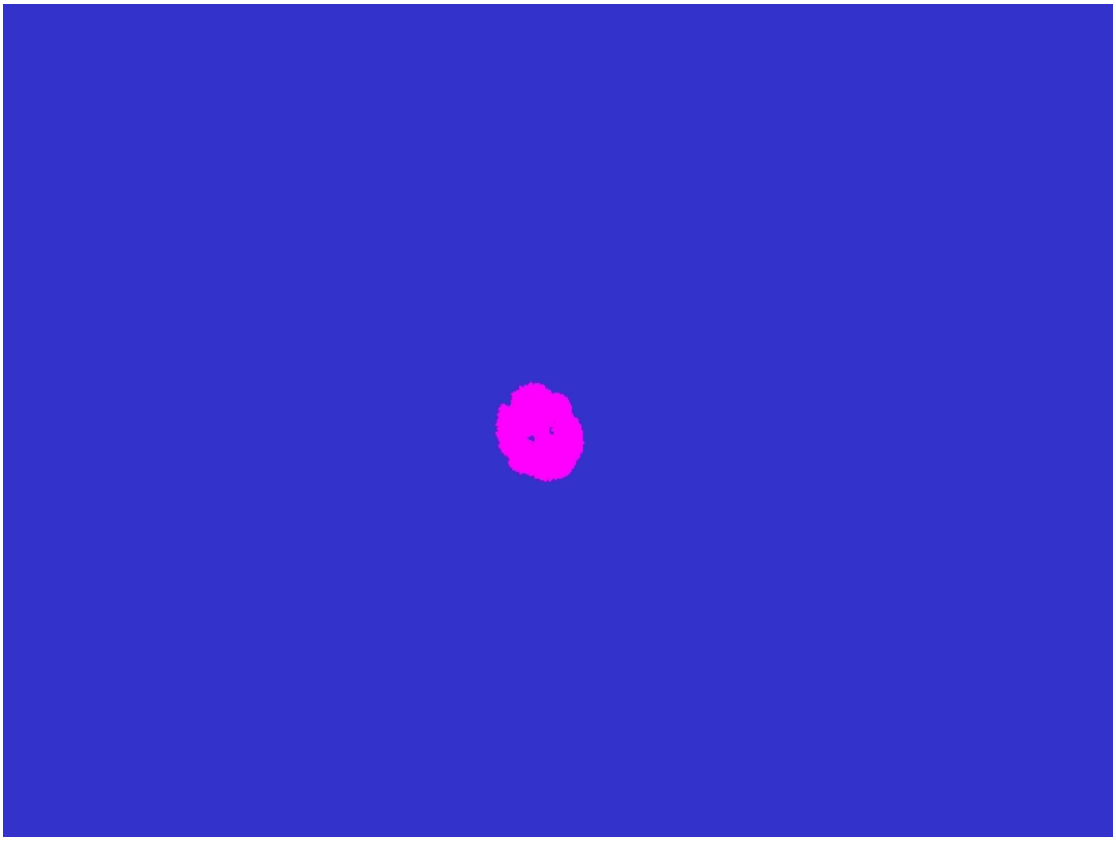


Figure 6.2.1.3 Pattern formed with range of adhesion force equals 100. All the agents gather and form one tight cluster

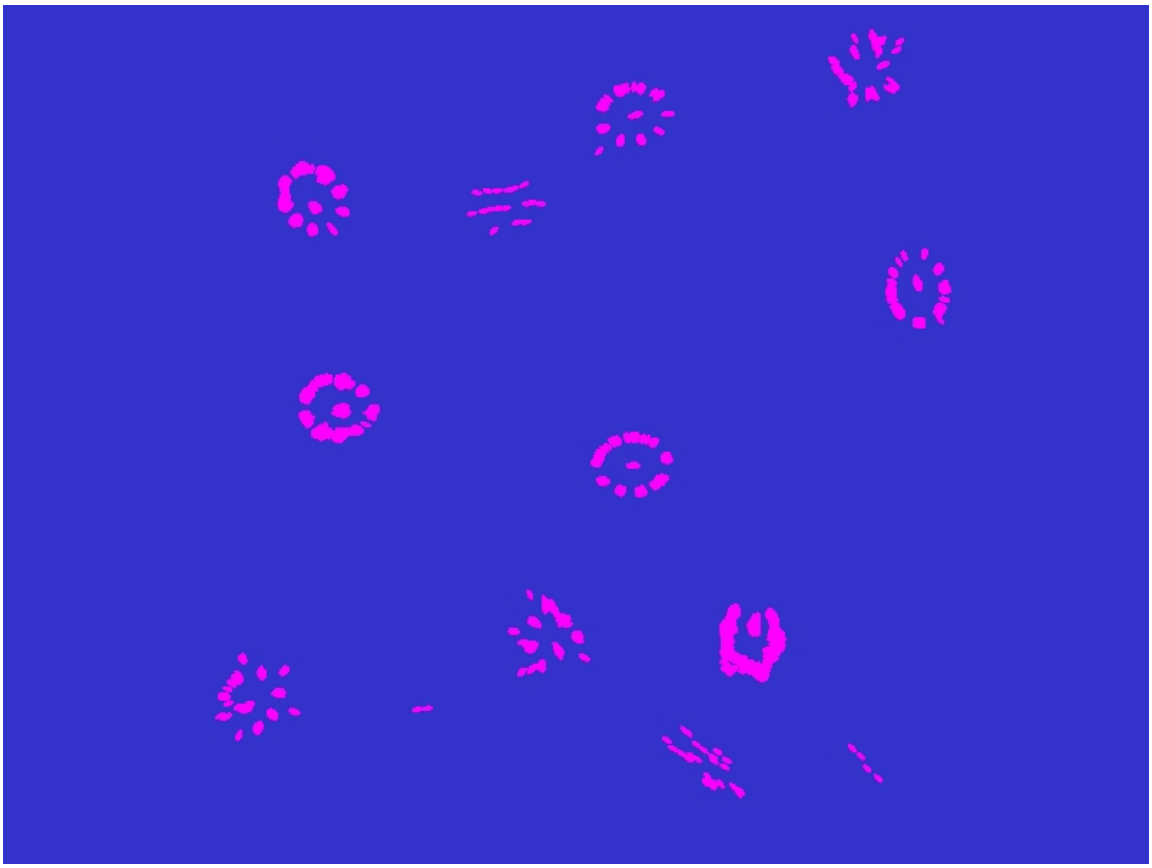


Figure 6.2.1.4 Pattern formed with range of adhesion force equals 10. Agents form

several clusters.

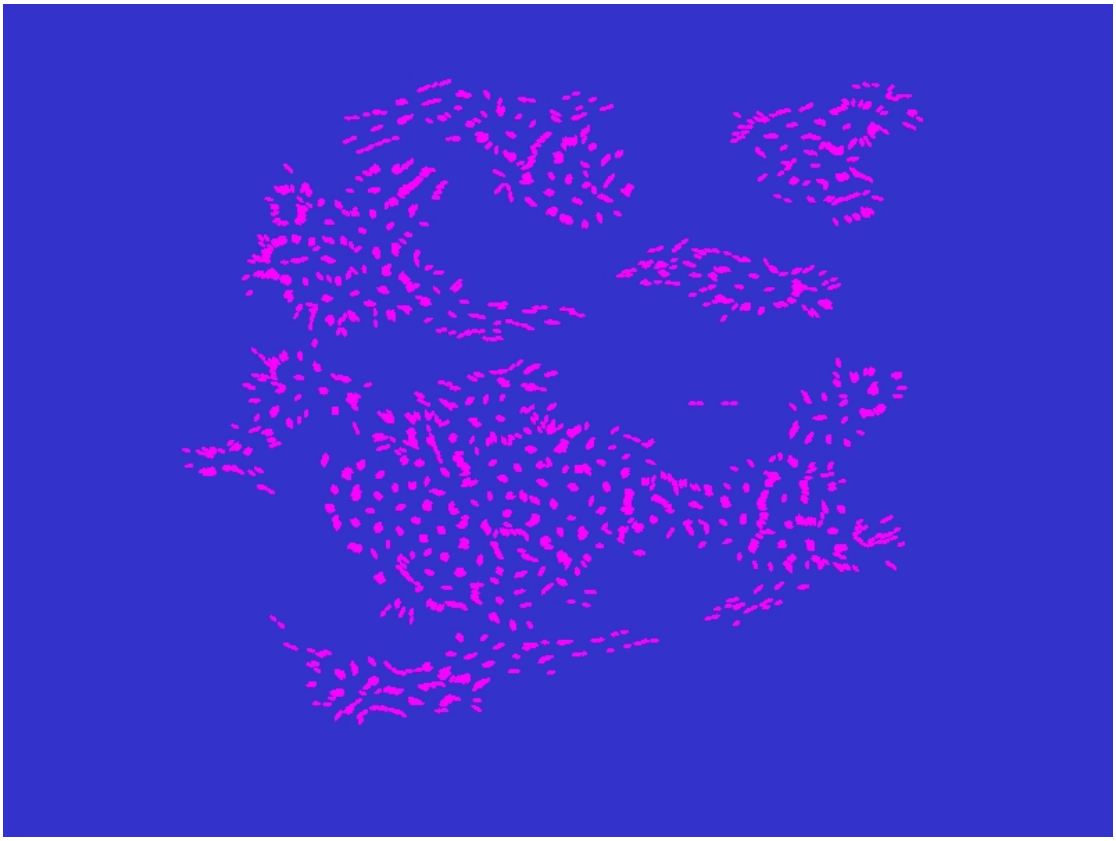


Figure 6.2.1.5 Pattern formed with range of adhesion force equals 3. Agents form more but smaller clusters.

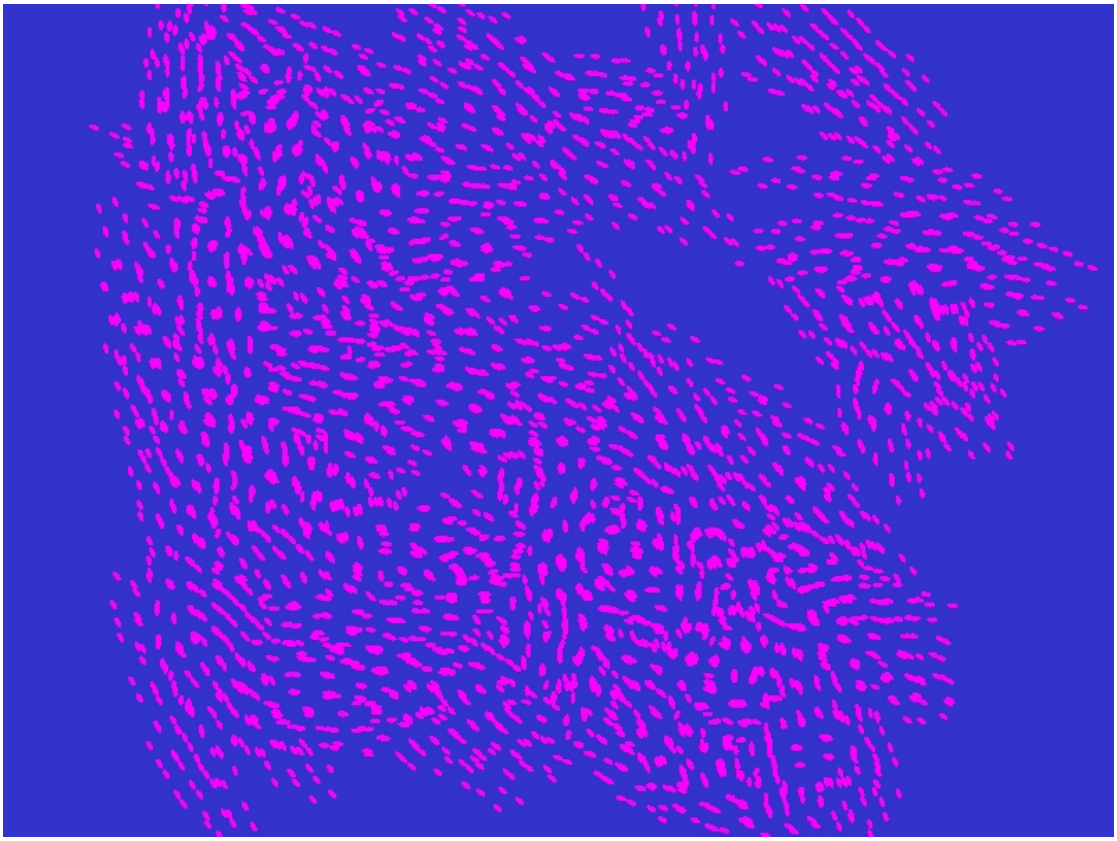


Figure 6.2.1.6 Pattern formed with range of adhesion force equals 1. Some agents form small clusters. Some agents are too far away from their neighbours to be attracted together and remain separated.

The corresponding pairwise correlation curves are shown in Figure 6.2.1.7.

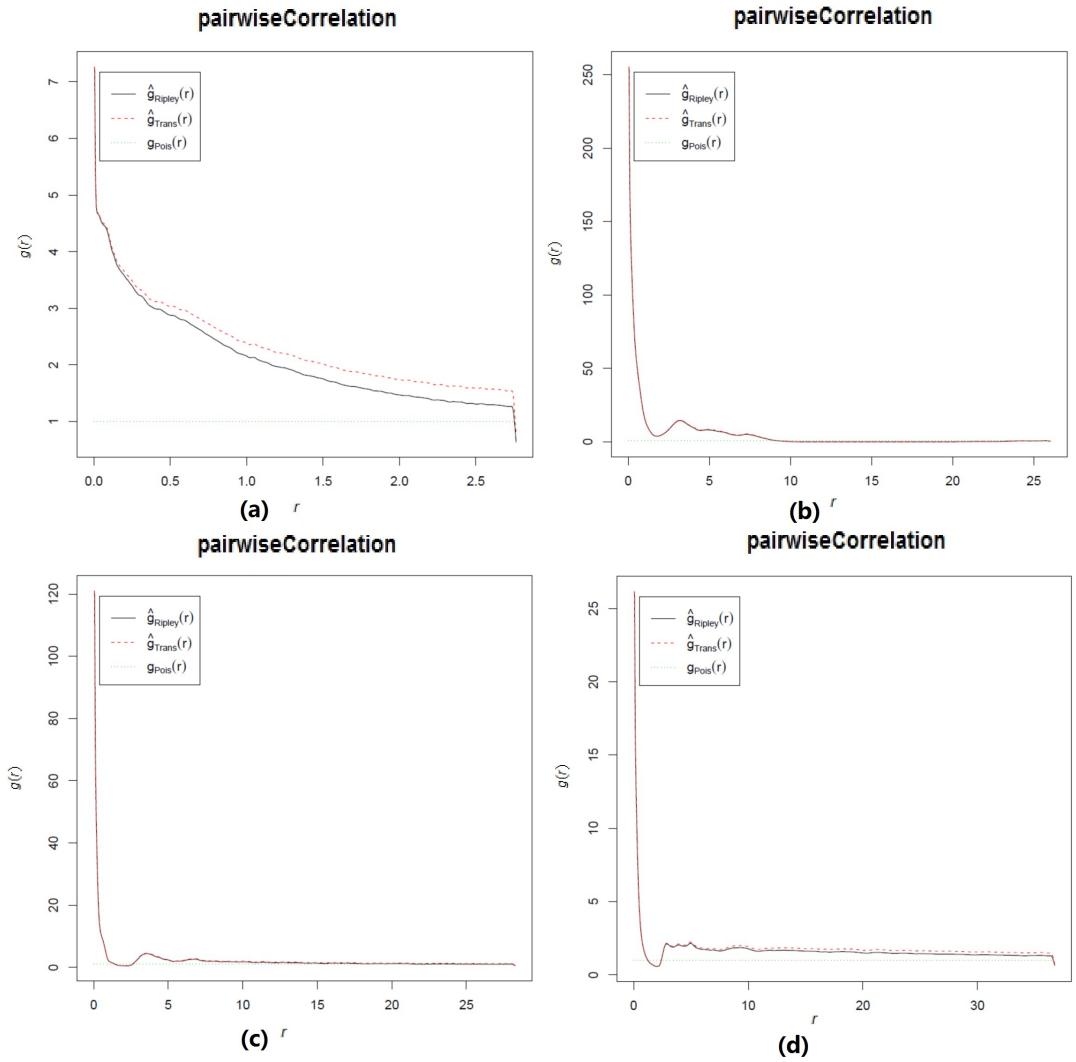



Figure 6.2.1.7 (a) Pairwise correlation curve with range of adhesion force equals 100; (b) pairwise correlation curve with range of adhesion force equals 10; (c) pairwise correlation curve with range of adhesion force equals 3; and (d) pairwise correlation curve with range of adhesion force equals 1.

The longer the effective range is, the more agents are stick together. Figure 6.2.1.7 (a) which is produced from Figure 6.2.1.3 shows a short distribution range (2.5), because the adhesion force acts on a long range, all the agents gather up together and form one big cluster; and on any distance within the distribution range, the curve is above $g(r)=1$ which means agents cluster on all distance. Figure 6.2.1.7 (b) is produced from Figure 6.2.1.4, in which the effective range of adhesion force is 10 and agents gather to form several clusters. Thus the distribution range in the curve is larger than

the top left curve. The curve forms a peak near distance=5 and a lowest point near distance=2.5, which means agents tend to gather up at distance=5 and disperse at distance=2.5. Figure 6.2.1.7 (c) is produced from Figure 6.2.1.5, in which the range of adhesion force is further decreased to 3. From the curve I can see that the typical distance between pair of agents remains the same, while the distribution range slightly increases. Figure 6.2.1.7 (d) is produced from Figure 6.2.1.6. The distribution range in this curve is larger (>30), because in this figure the range of adhesion force equals 1, only nearby agents may attract each other, and many agents are not close enough to be attracted together. The degree of gather up decreases from Figure 6.2.1.7 (b), Figure 6.2.1.7 (c) and Figure 6.2.1.7 (d), as in the corresponding figures, fewer agents gather to form clusters. The positions of peak and lowest point of curve are nearly the same, thus I consider it not controlled by the range of adhesion force.

## Point that adhesion force and contact force balance

The contact force acts when agents are in contact as a repulsive force. The more two agents overlap, the larger the contact force is. The direction of contact force is always opposite to direction of adhesion force. There is a point where 2 forces have same size and therefore balance, which controls the agents to remain static state. At static state the agents do not move closer to each other, nor move away from each other. The balancing point is represented by value of potential of agent at balancing point. The larger the value of potential is, the less overlap agents get when they are static, which represents less elasticity of agents. On the other hand, the smaller the value of potential is, the more overlap agents get, which represents more elasticity of agents.

In the explorative simulations to test the effect of balancing point, I keep the settings

for range of adhesion force = 1. Because larger range of adhesion force forms larger clusters from which it is hard to compare the difference among balancing point settings. I choose potential value = 0.9 to represent less elastic agents and potential value = 0.5 to represent more elastic agents. The patterns generated with the 2 settings are shown in Figure 6.2.1.8 and 6.2.1.9 separately.

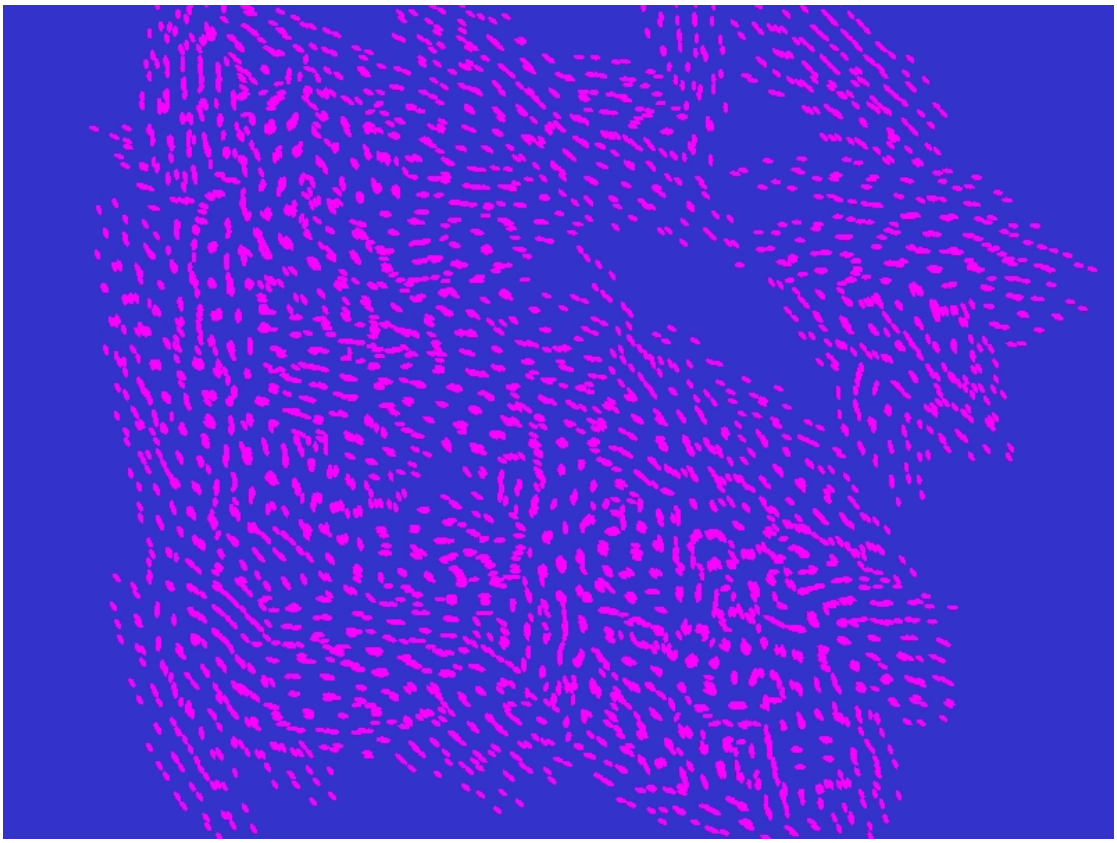


Figure 6.2.1.8 Balancing point: potential = 0.9; range of adhesion force equals 1

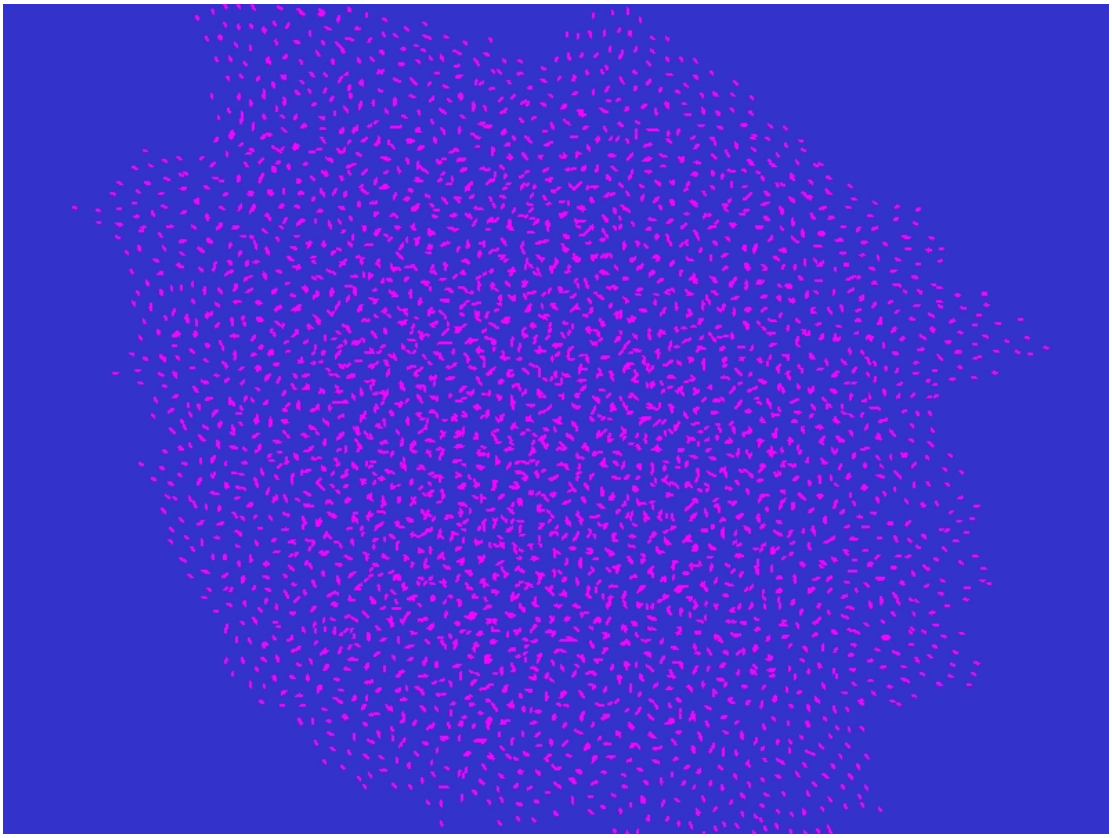


Figure 6.2.1.9 Potential = 0.5; range of adhesion force equals 1

Again I draw pairwise correlation curve to have a direct comparison of the 2 patterns, as shown in Figure 6.2.1.10.

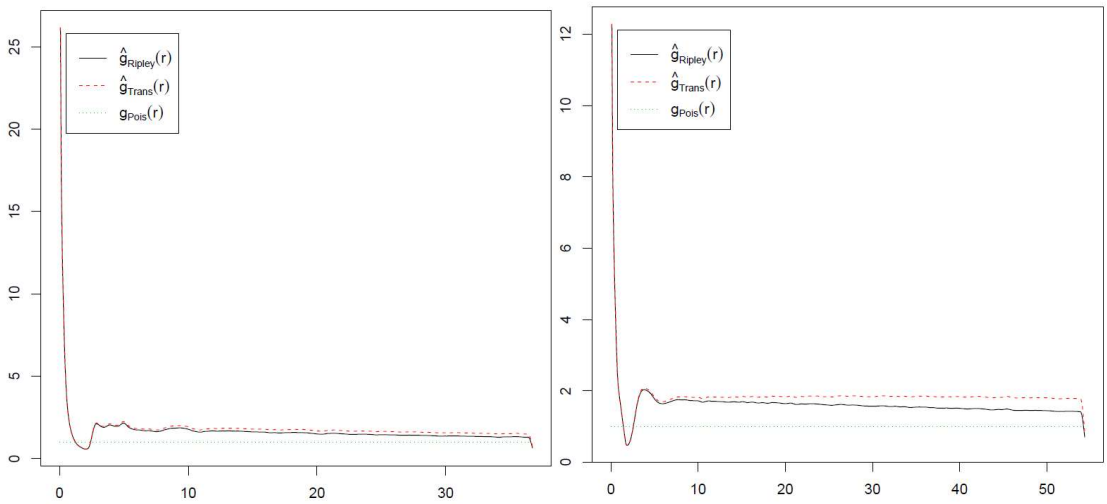



Figure 6.2.1.10 pairwise correlation curve of patterns with potential = 0.9 (left) and potential = 0.5 (right)

From Figure 6.2.1.10, I can see that when contact (repulsive) and adhesion (attractive)

force balance at high potential point, the degree of gathering value increases, and the distribution range decreases. However the position of lowest point and peak do not change. Therefore the position of lowest point and peak are not affect by balancing point either. I assume that they are only affected by shape of agents.

## Death rate

Although the death rate is studied and fitted in Chapter 5, its effect on pairwise correlation curve is explored in this section. In the following 3 simulation result, the agents have 80%, 40% and 20% chance to die when they finish cell cycle. The dead agents are removed from simulation.

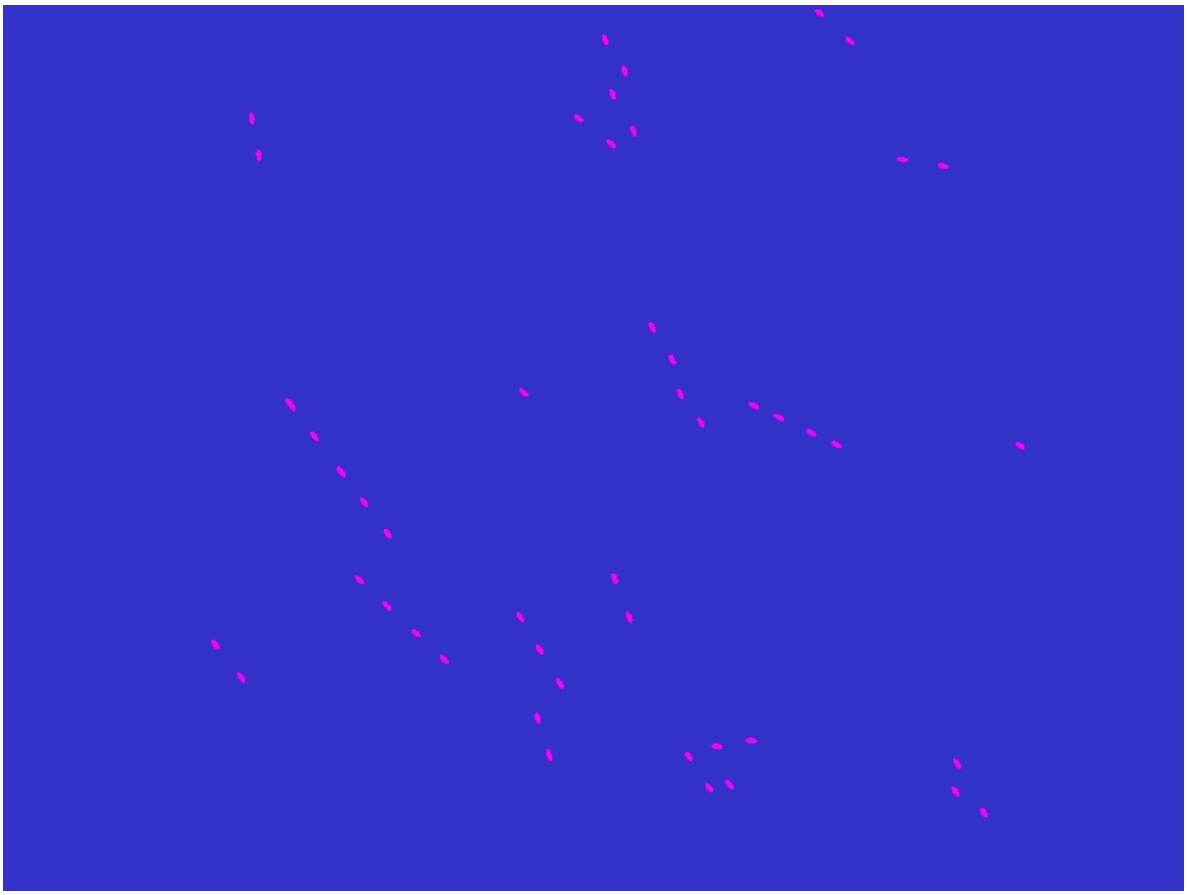


Figure 6.2.1.11 Pattern formed with death rate = 80%

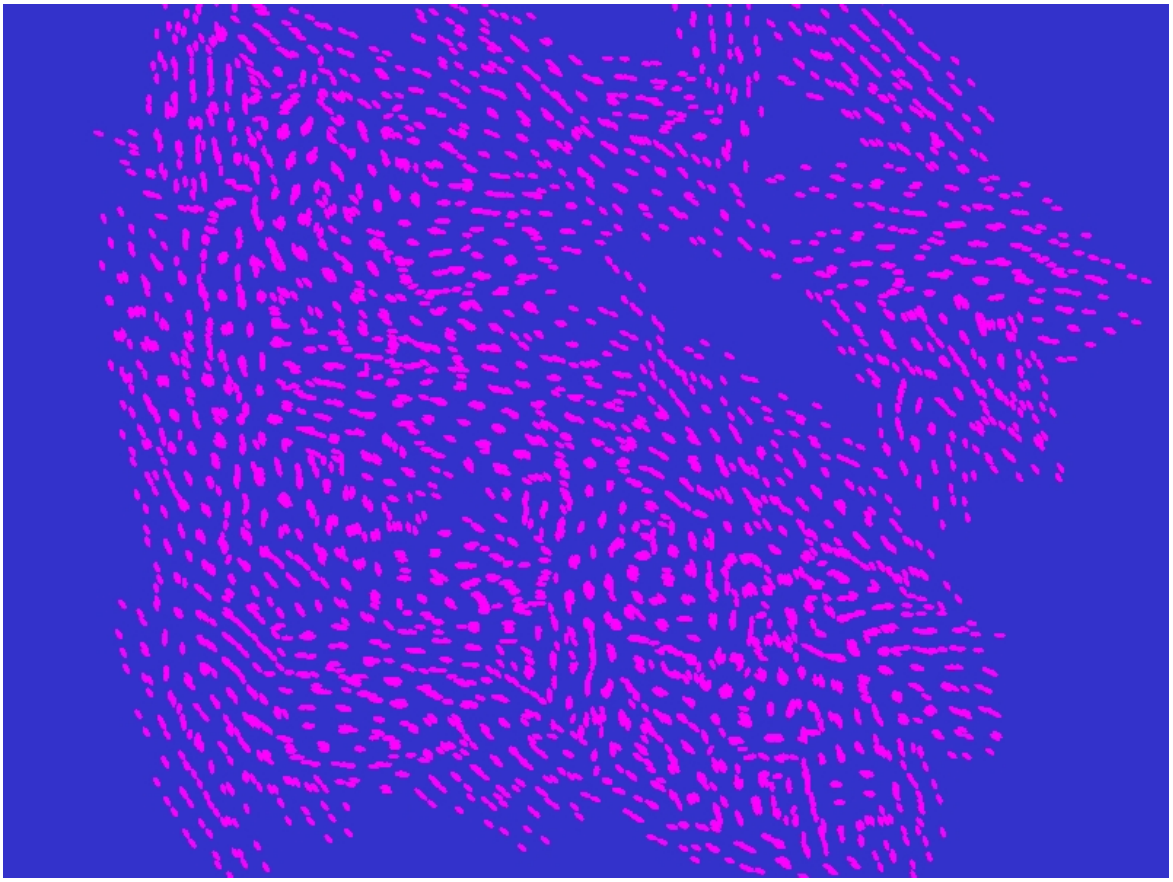


Figure 6.2.1.12 Pattern formed with death rate = 40%

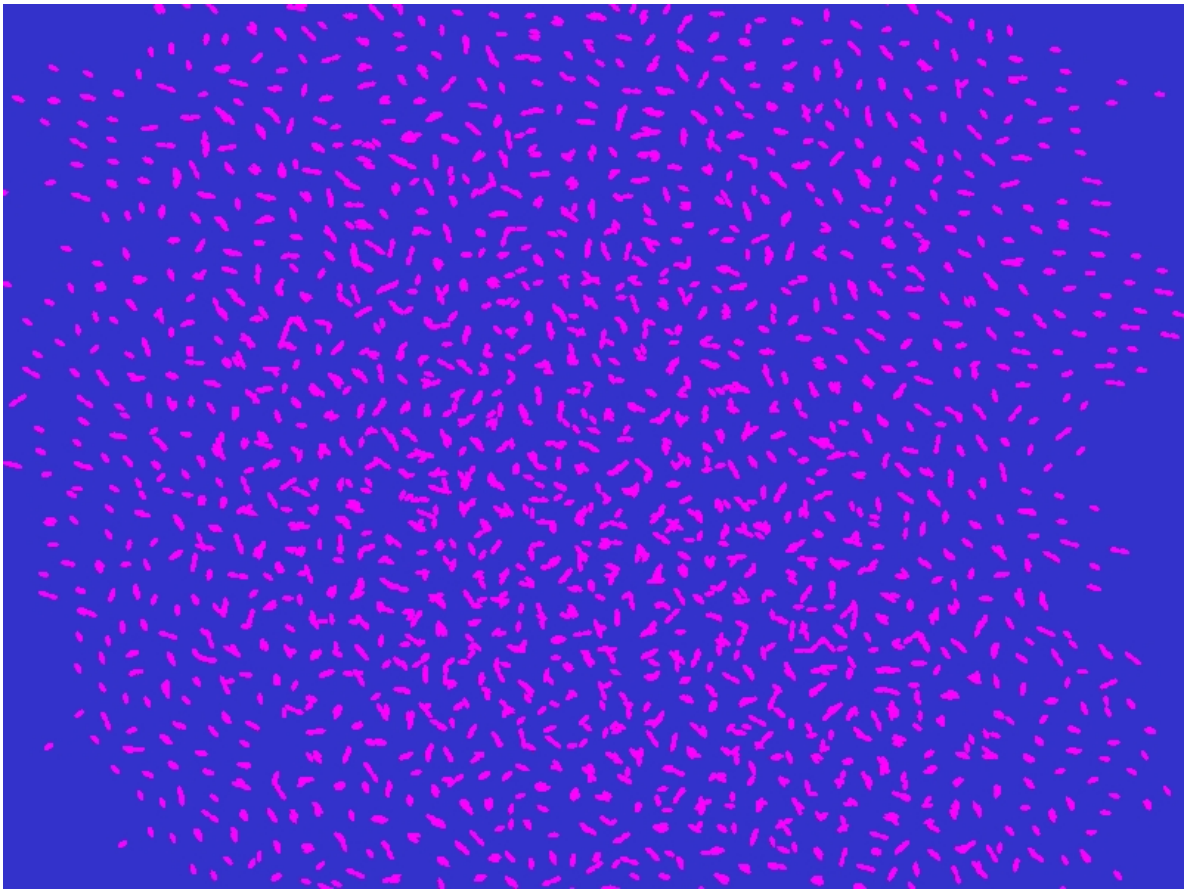


Figure 6.2.1.13 Pattern formed with death rate = 20%

The pairwise correlation curve produced from Figure 6.2.1.11 is shown in Figure 6.2.1.14 (a), the pairwise correlation curve produced from Figure 6.2.1.12 is shown in Figure 6.2.1.14 (n), and the pairwise correlation curve produced from Figure 6.2.1.13 is shown in Figure 6.2.1.14 (c).

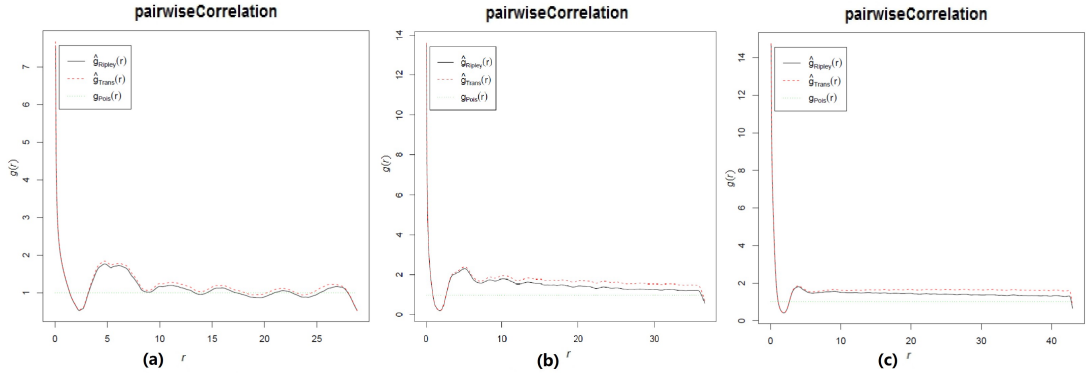



Figure 6.2.1.14 (a) Pairwise correlation curve of death rate = 80%; (b) pairwise correlation function of death rate = 40%; (c) pairwise correlation function of death rate = 20% (right).

From Figure 6.2.1.14, I can see that the higher the death rate is, the lower the degree of gathering is, and the shorter the distribution range is. This is reasonable, as with lower death rate there are more agents in the simulation result, thus for any agent there is higher chance to find another agent.

## 6.2.2 Fitting to the 5FU experiment

The time-lapse image and corresponding pairwise correlation function for the experimental data associated with the 5FU experiment is shown in Figure 6.2.2.1. According to Section 2.5.2, the shape of function shows that repulsion exists in the distribution of cancer cells in growth experiment with effect of 5FU. At the starting point of the experiment (Figure 6.2.2.1, time-lapse image (a) and pairwise correlation function (d)), there is a feature peak near distance = 20, which indicates the typical distance between typical cell and its neighbours. For distance >40 the cells tend to distribute randomly. However it is assumed that the agents are randomly distributed at the starting point, thus the experimental data does not match the assumption. The reason may be that, the time-lapse image covers a small part of one well on petri-dish, and this small area contains local structures. But the simulation will start from random distribution, otherwise it would be impossible for simulations of 5FU, hypoxia and combination group start from the same condition, and thus the difference of structure cannot be compared.

In the middle of experiment (Figure 6.2.2.1, time-lapse image (b) and pairwise correlation function (e)), the peak is also at distance=20 but with a reduced value,

which means the number of cells in clusters is reduced. There is a second peak around distance=35, which indicates that from a typical cell, there are groups of neighbours on this distance. Further, a valley appears at distance=60 with a value smaller than 1.0, and this means that cells tend to disperse at that distance.

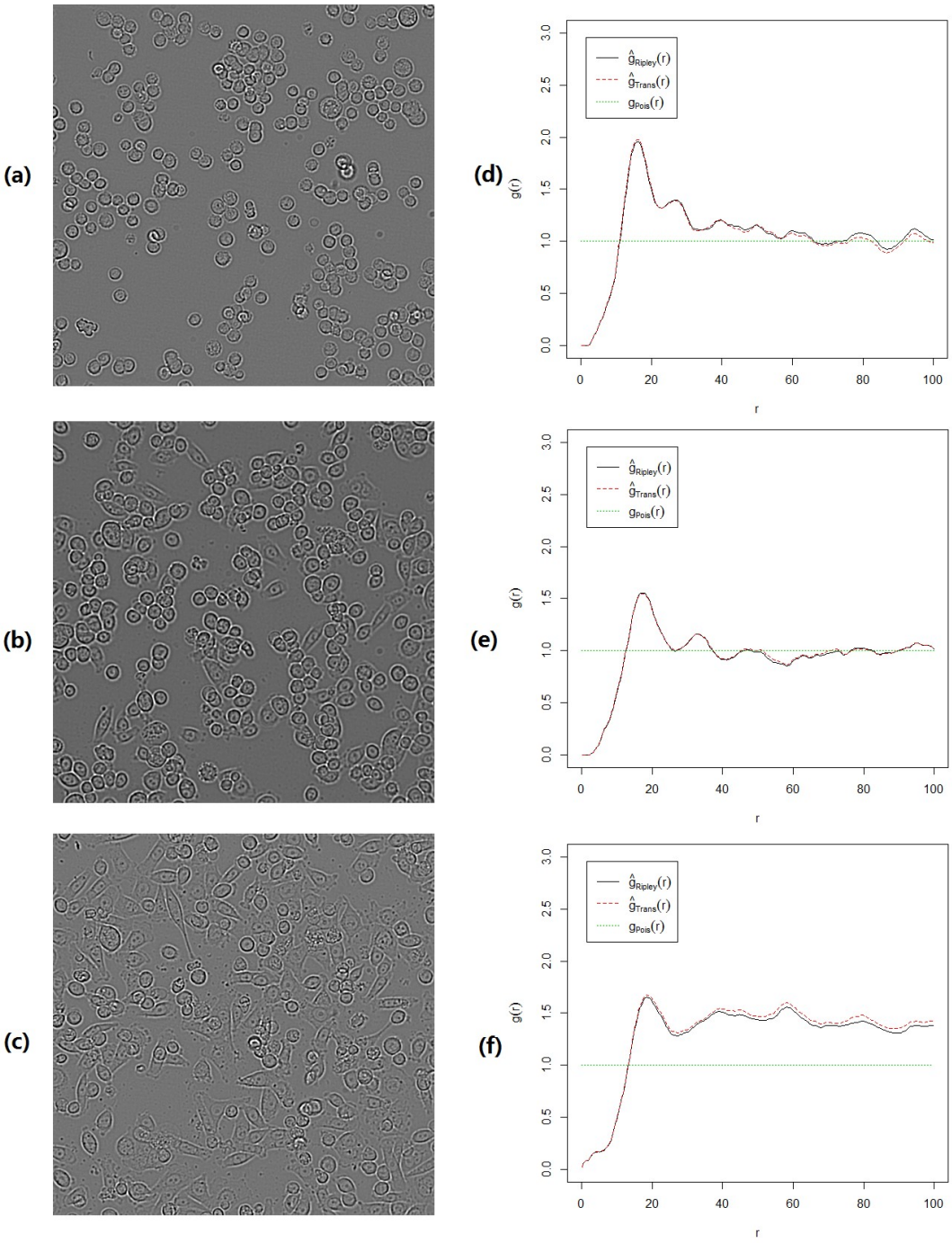



Figure 6.2.2.1 Time-lapse image and pairwise correlation curve of 5FU group of experiment. (a) is the time-lapse image of starting point of experiment, (d) is pairwise correlation

function of (a); (b) is the time-lapse image of middle point of experiment, (e) is pairwise correlation function of (b); (c) is the time-lapse image of the ending point of experiment, (f) is pairwise correlation function of (c).

At the end of experiment (Figure 6.2.2.1, time-lapse image (c) and pairwise correlation function (f)), the peak remains at the same distance. In addition, and in contrast to earlier time points, at all the distances greater than 20, the value of pairwise correlation is higher than 1.0, which means cells in experiment tend to gather on all distance. Figure 6.2.2.1 (c) shows that gaps exists, therefore the cells should not gather on all distance. However numerous cells are distributed on Figure 6.2.2.1 (c) with relative short distance, which is identified as one large group of cells. To display the spatial feature correctly, a circular observation window has to be applied to Figure 6.2.2.1 (c), as shown in Figure 6.2.2.2. Note that because a circular observation window is chosen, the valid range of distance in pairwise correlation function is limited.

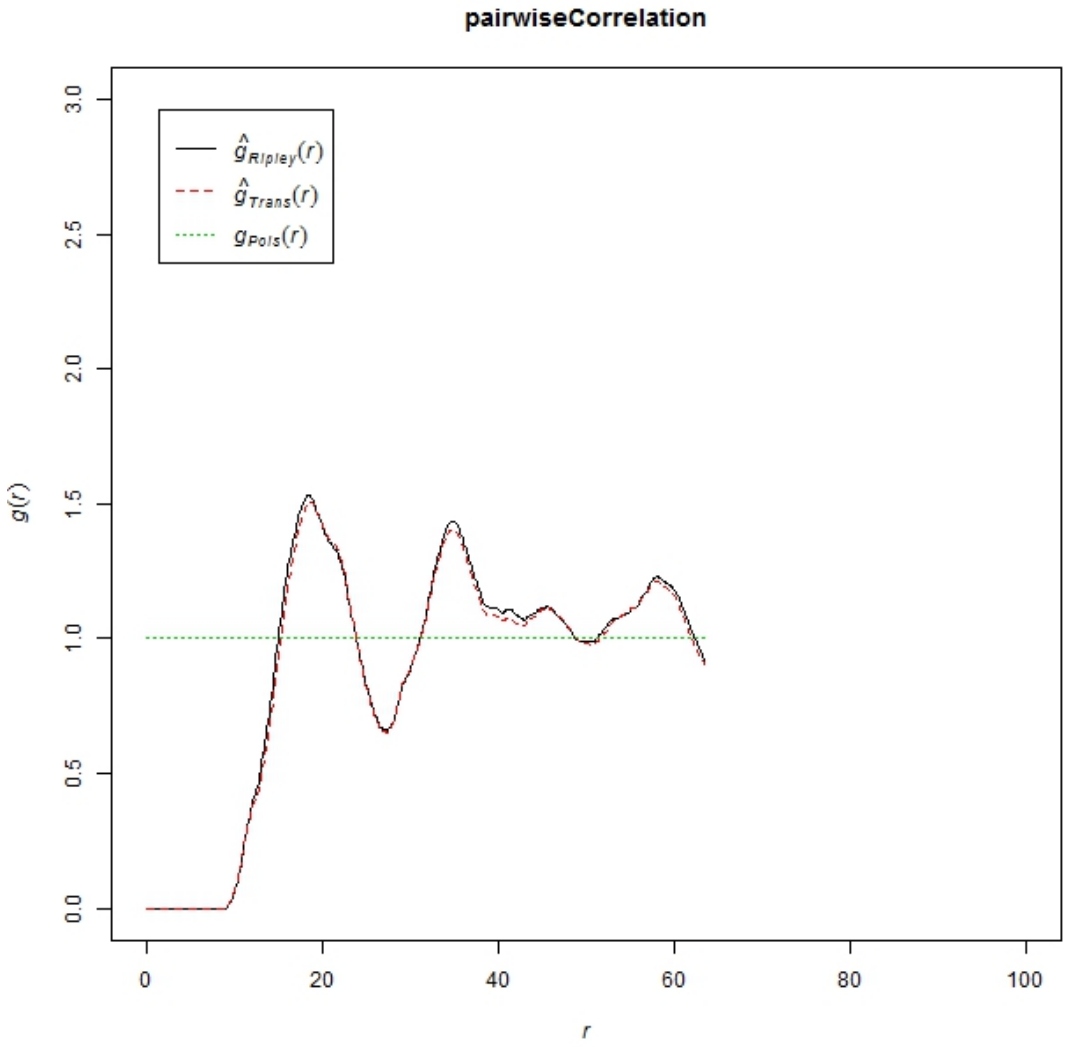



Figure 6.2.2.2 Pairwise correlation function of ending point of experiment with 5FU, observed with a circular window. The range of distance is limited because of the circular window.

Although the valid range of distance in pairwise correlation function is limited, it displays more detail within the range. Comparing to Figure 6.2.2.1 (d) and (e), Figure 6.2.2.2 shows a deeper valley near distance=30, and the second peak is more obvious. In addition, there is a third peak near distance=60. In this way Figure 6.2.2.2 indicates a progress of cell cluster, which matches Figure 6.2.2.1 (c).

With the same population growth setting for the 5FU experiment as discussed in chapter 5, I run the simulation. Firstly I set the population in simulation so that the

number density of simulation is same with experiment, which are shown in Table 6.2.2.1. Therefore the simulation does not start with the same population of experiment, but with the same number density.

| | number density |
|---|---|
| Experiment group: 5FU, starting point | 0.00099 |
| Simulation: 5-FU, starting point | 0.00094 |
| Experiment group: 5FU, end point | 0.00054 |
| Simulation: 5-FU, end point | 0.00053 |

Table 6.2.2.1 The number density of experimental data and simulation data

A set of simulation result is shown in Figure 6.2.2.3. From (a) - (c), it is shown that agents are more and more clustered.

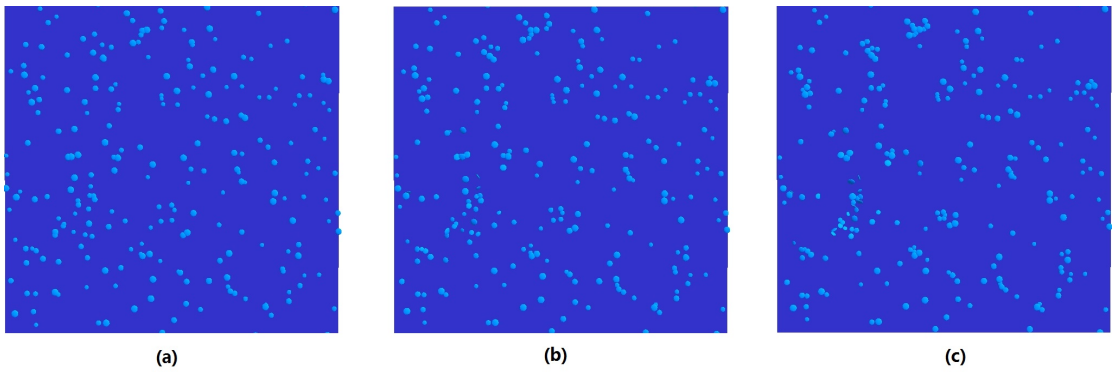



Figure 6.2.2.3 Simulation result of agents under effect of 5FU. The light blue part is the agents, while the dark blue part is background. (a) is the start point of simulation, (b) is the middle point of simulation, and (c) is the end point of simulation.

The pairwise correlation function of simulation data is also estimated. Unlike the experimental data, the first peak of simulation data is near distance=0, while the first peak of experimental data is observed near distance=20, as shown in Figure 6.2.2.4.

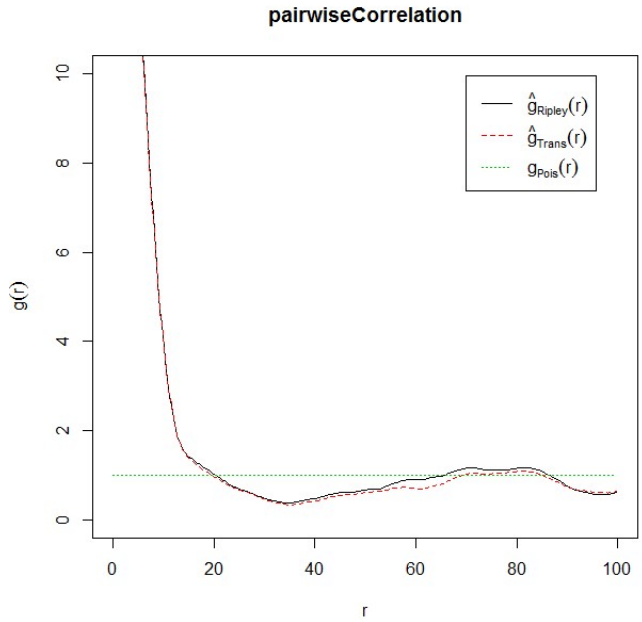



Figure 6.2.2.4 End point of simulation of agents with effect of 5FU. A high peak appears near distance=0.

According to the interpretation of the pairwise correlation function, values greater than 1.0 represent clustering. Analysing the position data of agents in the simulation, I find agents that are closer than 2 units in distance. However, agents should not be so close together. One possible reason is that some agents 'climb on top of' other agents in the simulation. In the cell growth experiments, because the time-lapse images are taken with top-down view, a similar phenomenon cannot be observed. Thus a filter is introduced to process the simulation data so that agents that are too are discarded. The processed simulation data does not produce the high peak near distance=0, but there is a side effect, which is that the uniform density of simulation data is reduced and then is lower than experimental value. The pairwise correlation function of filtered simulation data is shown in Figure 6.2.2.5, in which the first peak appears near distance=10.

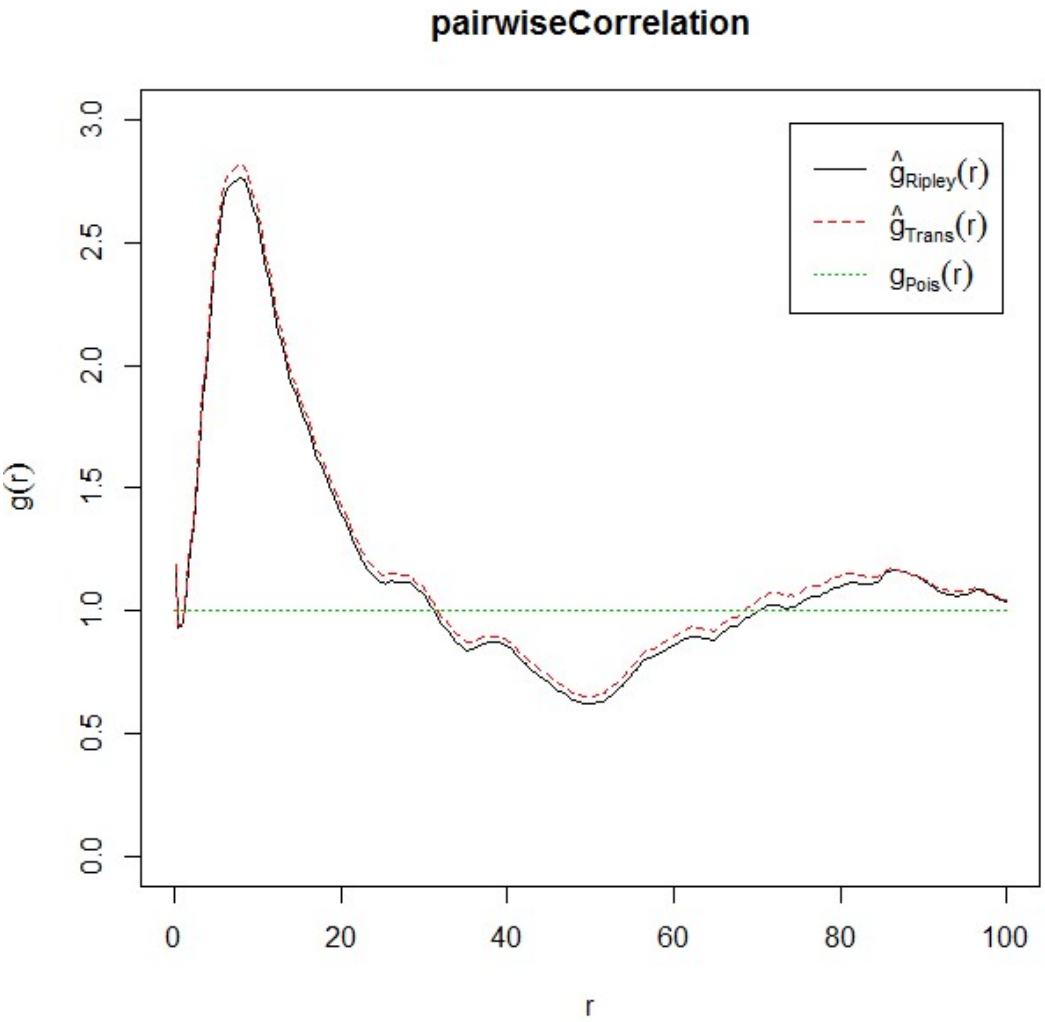



Figure 6.2.2.5 Pairwise correlation curve of filtered simulation data. The first peak appears near distance=10.

Running the 5FU simulation 10 times with this filter, I can draw the average pairwise correlation function, as Figure 6.2.2.6 shows. Note that the range of distance in Figure 6.2.2.6 (c) is reduced because this pairwise correlation function is generated from a circular observation window.

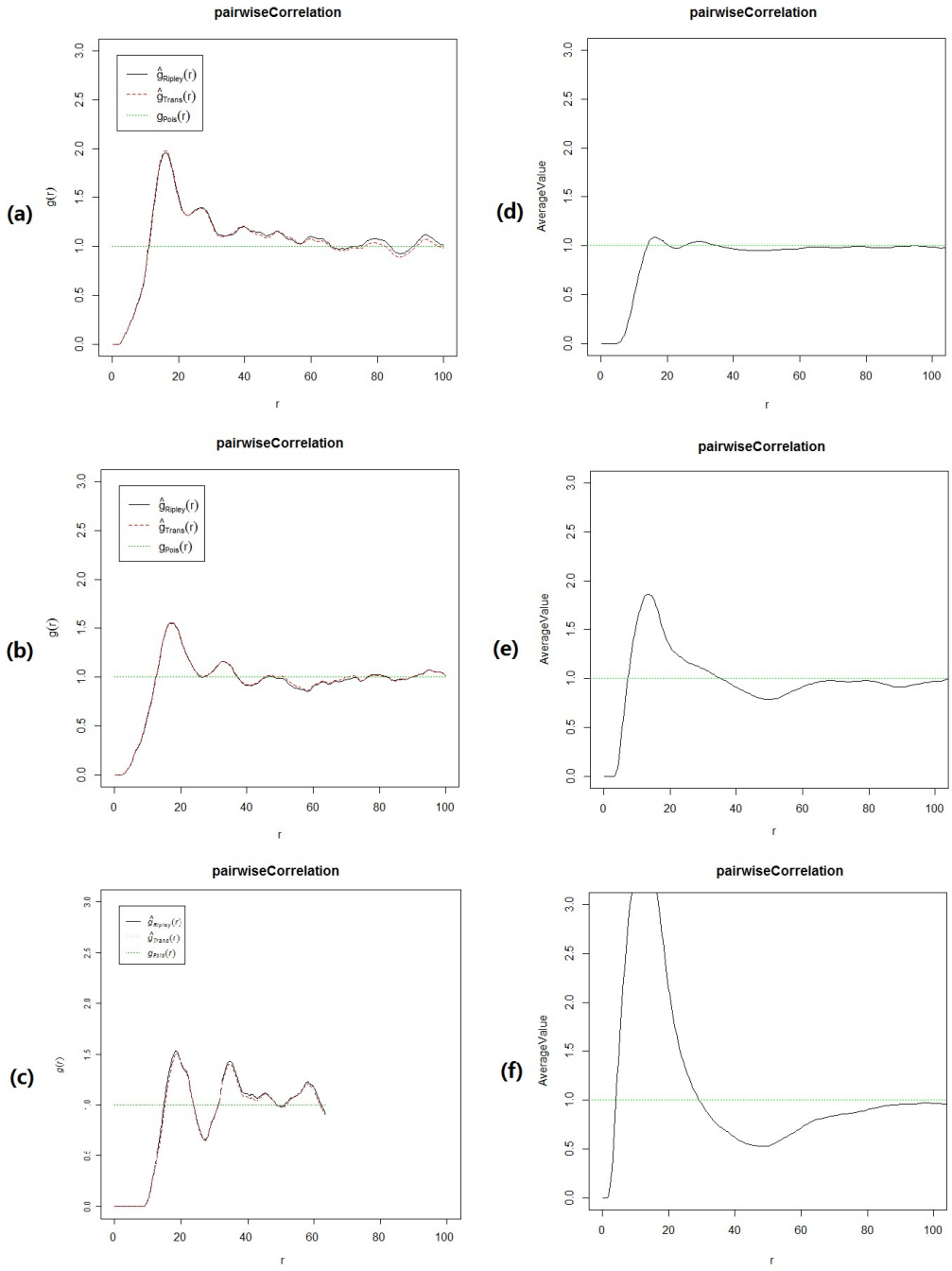



Figure 6.2.2.6 Pairwise correlation function of experimental data (a) - (c), and average pairwise correlation function of 10 simulations (d) - (f) of 5FU group. (a) and (d) are the starting point of experimental and simulation data; (b) and (e) are middle point of experimental and simulation data; and (c) and (f) are end point of experimental and simulation data.

From Figure 6.2.2.6 (d), it is observed that, the simulation starts from near-random distribution. I managed to reproduce a similar distribution in simulation in the middle

point with the one in experiment, as shown in Figure 6.2.2.6 (b) and (e), both pairwise correlation functions have a peak near distance=20 and a valley near distance=60; plus the cells and agents distribute randomly on distance>60. From Figure 6.2.2.6 (d) to Figure 6.2.2.6 (e), more agents gathers in cluster, thus in Figure 6.2.2.6 (f) this trend continues and results larger clusters. Note that to keep the same ratio of x and y-axis of Figure 6.2.2.6 (f) with other figures, the peak of pairwise correlation function exceeds the range of y-axis.

As shown in Figure 6.2.2.6 (a), the experiment does not start from random distribution, it is a different process from Figure 6.2.2.6 (a) to Figure 6.2.2.6 (b) with Figure 6.2.2.6 (d) to Figure 6.2.2.6 (e), which is the reason that the end point of simulation (Figure 6.2.2.6 (f)) has different spatial feature from experimental data (Figure 6.2.2.6 (c)).

## 6.2.3 Fitting to the hypoxia experiment

The time-lapse images of starting, middle and end point of experiment under effect of hypoxia is shown in Figure 6.2.3.1 (a) - (c). The pairwise correlation function of Figure 6.2.3.1 (a) is shown in Figure 6.2.3.1 (d); the pairwise correlation function of Figure 6.2.3.1 (b) is shown in Figure 6.2.3.1 (e); and the pairwise correlation function of Figure 6.2.3.1 (c) is shown in Figure 6.2.3.1 (f).

As shown in Figure 6.2.3.1 (d), the experiment of hypoxia group does not start with random distribution either, like the case in 5FU group, but a repulsive distribution with the first peak near distance=20. Then from Figure 6.2.3.1 (d) to Figure 6.2.3.1 (e), the $g(r)$ value on the first peak is reduced, which means that there are less cells in the

clusters. From Figure 6.2.3.1 (e) to Figure 6.2.3.1 (f), the $g(r)$ value of the peak is slightly increases and the part of peak above the random-distribution-line (green dotted line) is also increased, which suggests that more cells gathers and form less tight clusters.

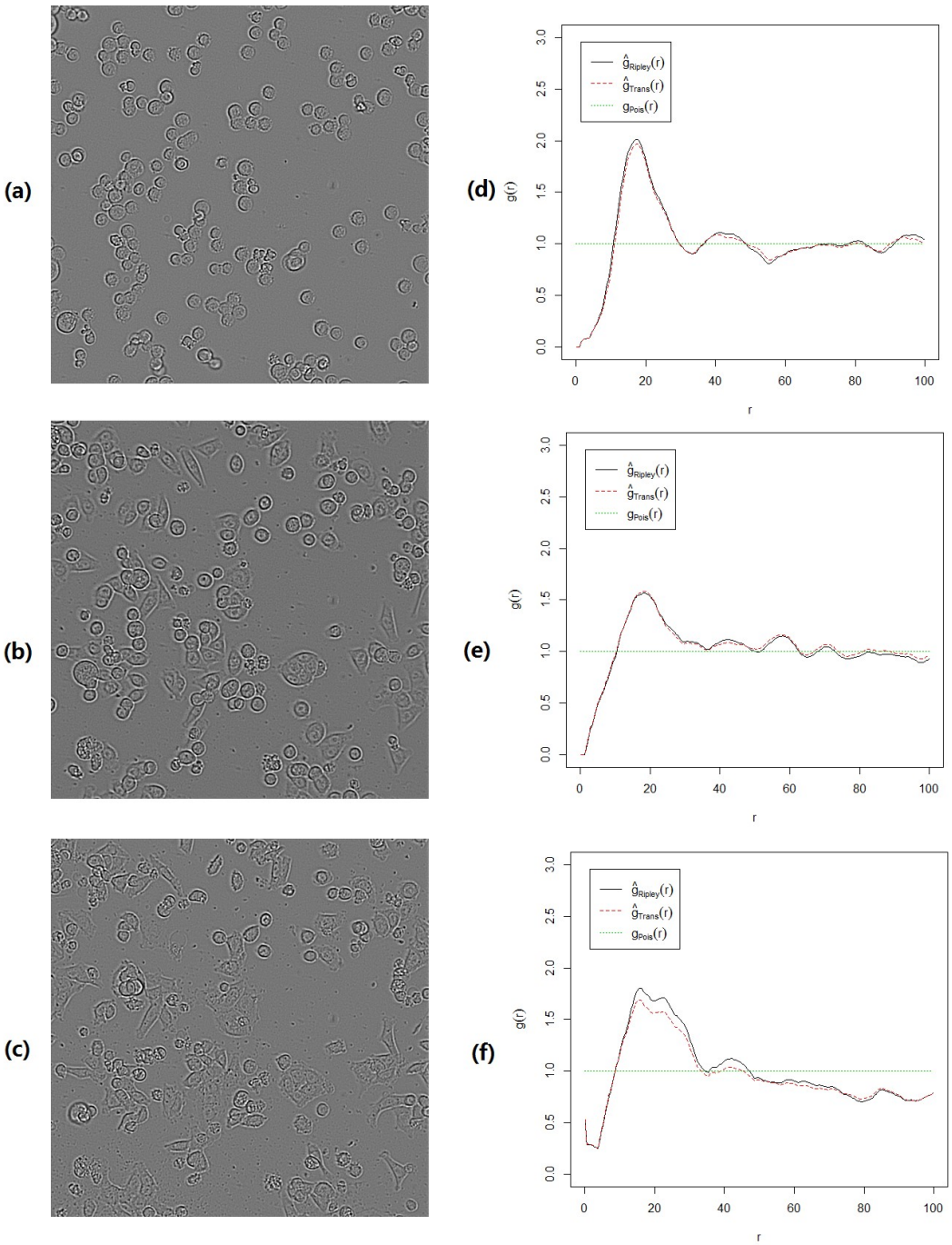



Figure 6.2.3.1 Time-lapse image and pairwise correlation curve of hypoxia group of

experiment. (a) is the time-lapse image of starting point of experiment, (d) is pairwise correlation function of (a); (b) is the time-lapse image of middle point of experiment, (e) is pairwise correlation function of (b); (c) is the time-lapse image of the ending point of experiment, (f) is pairwise correlation function of (c).

The simulation of cell growth and patterns in a hypoxic condition begins with similar number density of the experiment, as shown in Table 6.2.3.1. According to our biological experts, the cells under the effect of hypoxia change their metabolism, which results slower motility. Thus in the simulation, and by systematic variation of parameters, the speed of agent is reduced to a quarter of speed of the 5FU simulation.

| | number density |
|---|---|
| Hypoxia experiment (start) | 0.00061 |
| Hypoxia simulation (start) | 0.00051 |

Table 6.2.3.1 The number density of experiment and simulation of hypoxia group. Note at the starting point, the value for simulation is lower because some agents are filtered due to close distance.

The simulation results are shown in Figure 6.2.3.2. As discussed in Section 6.2.2, the simulation starts from random distribution, as shown in Figure 6.2.3.2 (a). The red and green dots are two types of agents, as discussed in Section 5.5.3. Following the growth curve in Section 5.5.3, there are less agents in middle point of simulation (Figure 6.2.3.2 (b)) than the starting point of simulation (Figure 6.2.3.2 (a)), and less agents in the end point (Figure 6.2.3.2 (c)) than the middle point of simulation (Figure 6.2.3.2 (b)).

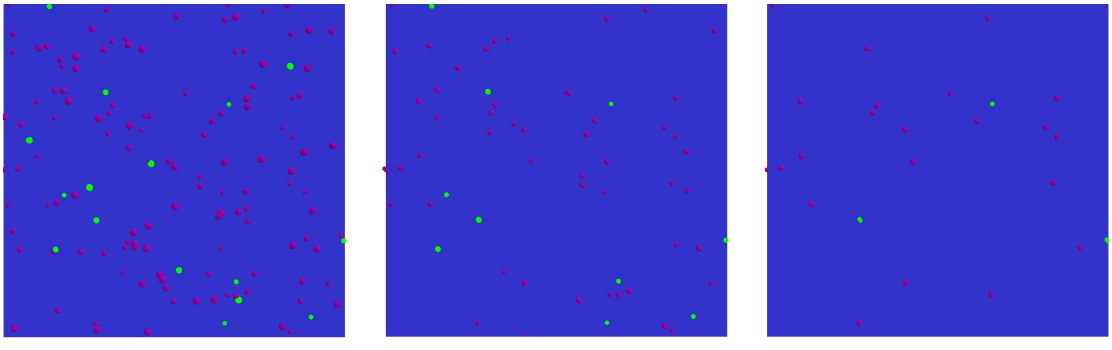


Figure 6.2.3.2. Simulation result of agents under effect of hypoxia. The red and green part are two types of agents, while the dark blue part is background. (a) is the start point of simulation, (b) is the middle point of simulation, and (c) is the end point of simulation.

To avoid the bias, 10 simulations are ran, and the average pairwise correlation function is shown in Figure 6.2.3.3 together with experimental function. The Figure 6.2.3.3 (a), (b) and (c) are the pairwise correlation function of starting, middle and end point of experimental data, while the Figure 6.2.3.3 (d), (e) and (f) are the pairwise correlation function of starting, middle and end point of simulation data. Comparing Figure 6.2.3.3 (a) and (d), it is observed that unlike the experimental data, the simulation starts from a random distribution. Then although Figure 6.2.3.3 (e) is not in the same shape of Figure 6.2.3.3 (b), Figure 6.2.3.3 (e) shows a less tight distribution than Figure 6.2.3.3 (d), which is the same trend from Figure 6.2.3.3 (a) to Figure 6.2.3.3 (b).

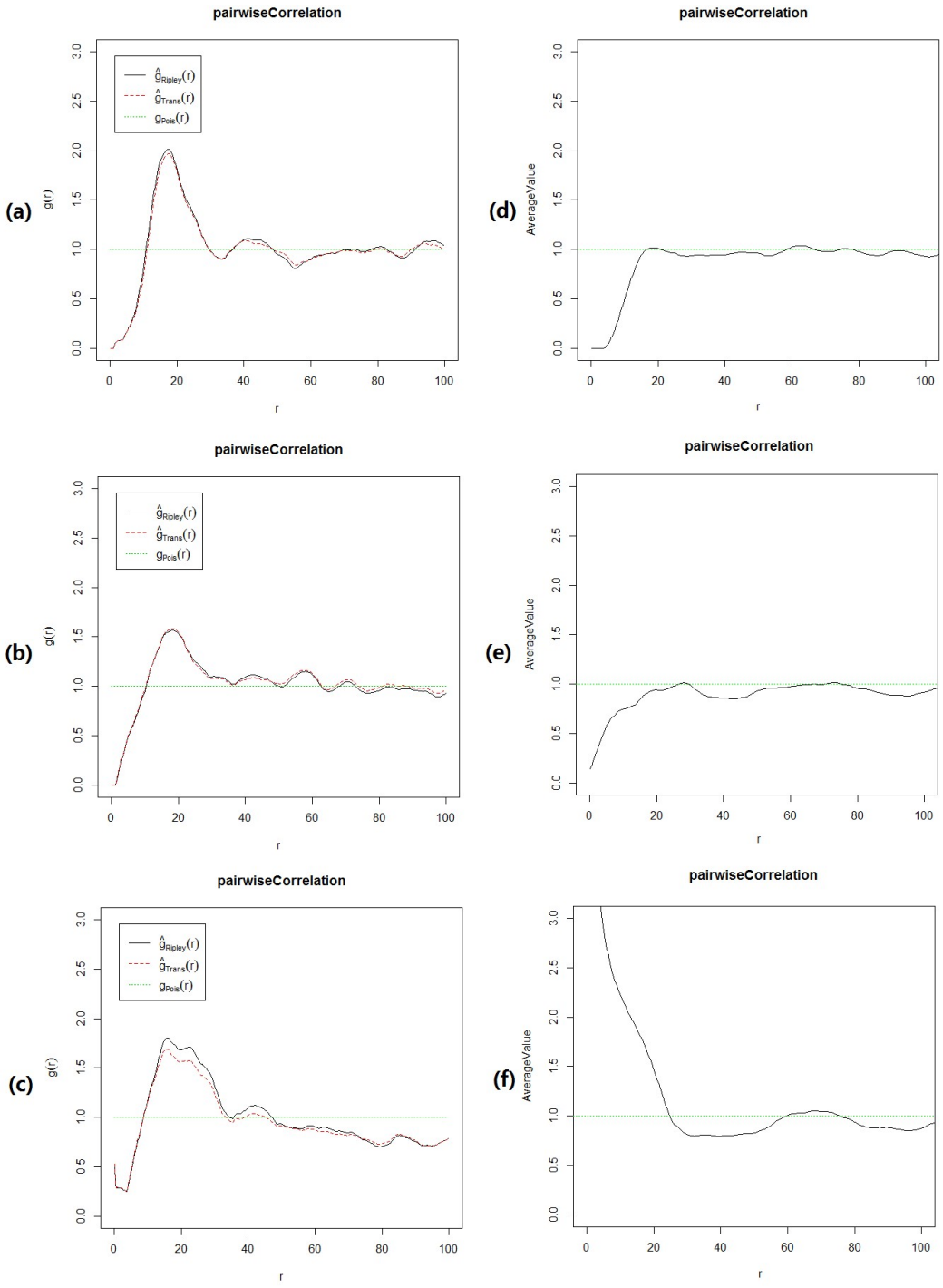



Figure 6.2.3.3 Pairwise correlation function of the starting, middle and end point of experimental data (a) - (c) and average of 10 simulations of hypoxia (d) - (f).

The end point of simulation generates a different pairwise correlation function from the end point of experiment. The difference between Figure 6.2.3.3 (b) and (e), and between Figure 6.2.3.3 (c) and (f), may be the difference of number density of cells

(agents) in experimental and simulation data, as shown in Table 6.2.3.2: the number density in simulation is significantly smaller than in experiment. As the growth curve in simulation of hypoxia group is calibrated with experimental data, and simulation and experimental data are close in Day 0 and Day 1, within the time range of time-lapse image (20 hours), the simulation and experimental data of number density of cell should be similar. This problem also may caused by the small area that the time-lapse image covers, in which the number density of cells is biased.

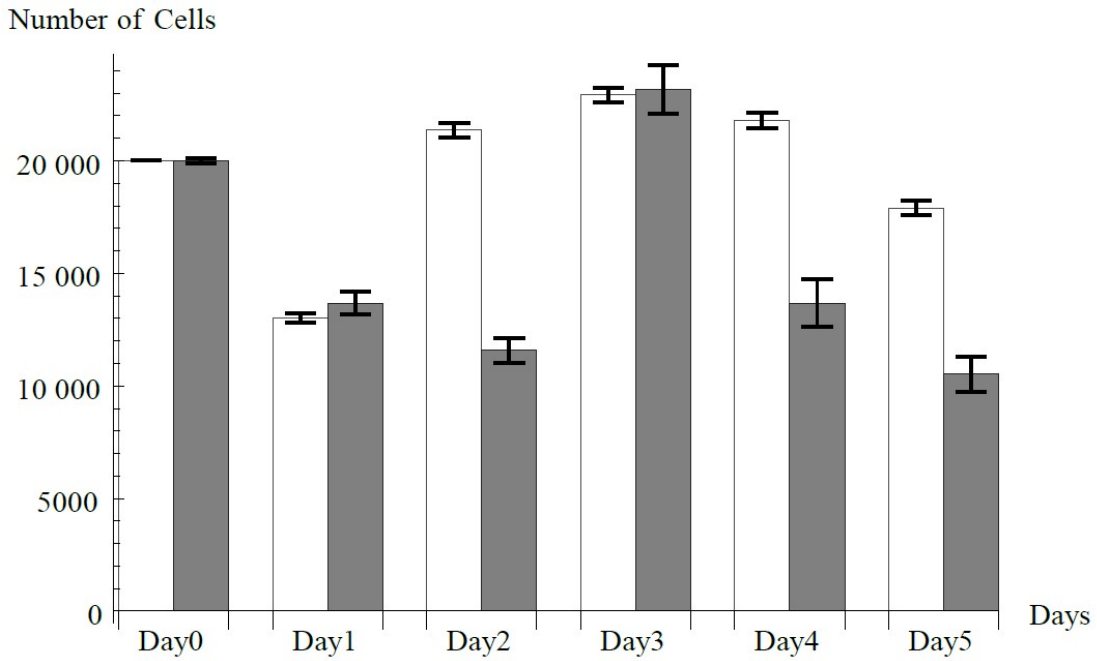



Figure 6.2.3.4 Hypoxia growth curve of my model (grey bars) with two types of agents and experiment result (white bars), in which the simulation and experimental data are very close in Day 0 and Day 1.

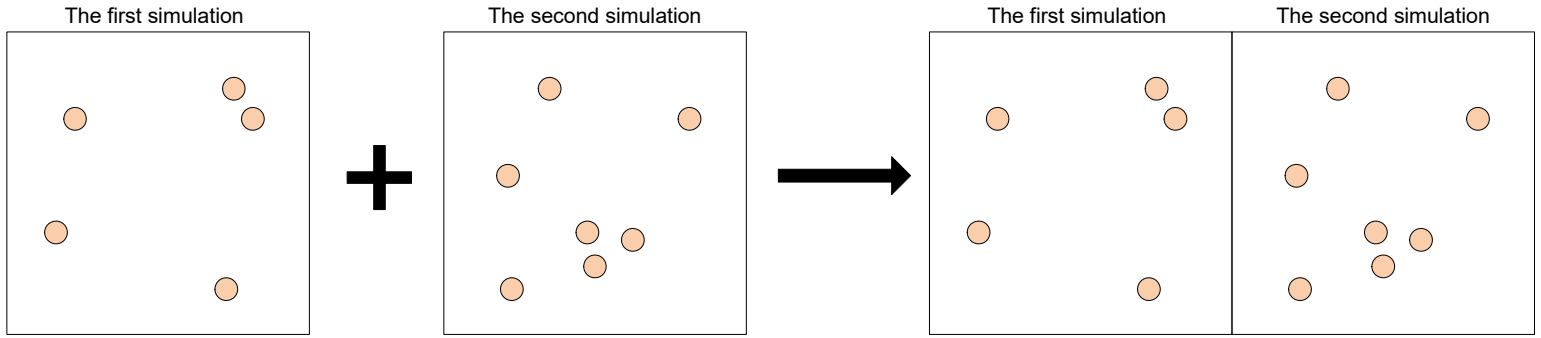
| | number density |
|---|---|
| Hypoxia group experiment: start point | 0.00061 |
| Hypoxia group simulation: start point | 0.00051 |
| Hypoxia group experiment: middle point | 0.00062 |
| Hypoxia group simulation: middle point | 0.00023 |
| Hypoxia group experiment: end point | 0.00052 |
| Hypoxia group simulation: end point | 0.00012 |

Table 6.2.3.2 Number density of cells (agents) in the starting, middle and end point of expeirment and simulation. At the starting point, the number density is similar in experiment and simulation, while it is significantly reduced in simulation data in the middle and end point, which does not match the calibration result of growth curve in

Figure 6.2.3.4.

Apart from the difference in number density between experimental and simulation data, the pairwise correlation function of the end point of simulation (Figure 6.2.3.3 (f)) shows a different type of distribution with experimental data (Figure 6.2.3.3 (c)). Plus Figure 6.2.3.3 (f) is not consistent with Figure 6.2.3.3 (d) and (e). As all agents that are closer than 10 units distance are already filtered, there should be no peak before distance=10. I assume this problem is caused by the low total number of the agents.

To test this assumption, 10 simulation results are combined together to generate a new pairwise correlation curve. The method is, keep the first simulation result unchanged; then add the x-coordinate of the second simulation result with width of simulation output so that the second simulation result is considered attached to the first one head to tail, as Figure 6.2.3.2 (top) shows. Similarly, add the x-coordinate of the third simulation result with double width of the simulation output so that the third simulation result is attached with the second one head to tail. Do the same thing to the rest 7 simulation results, then all 10 simulation results are attached head to tail and can be considered as one result with plenty of agents, as Figure 6.2.3.5 (bottom) shows. The pairwise correlation curve of the combined data is shown in Figure 6.2.3.6.



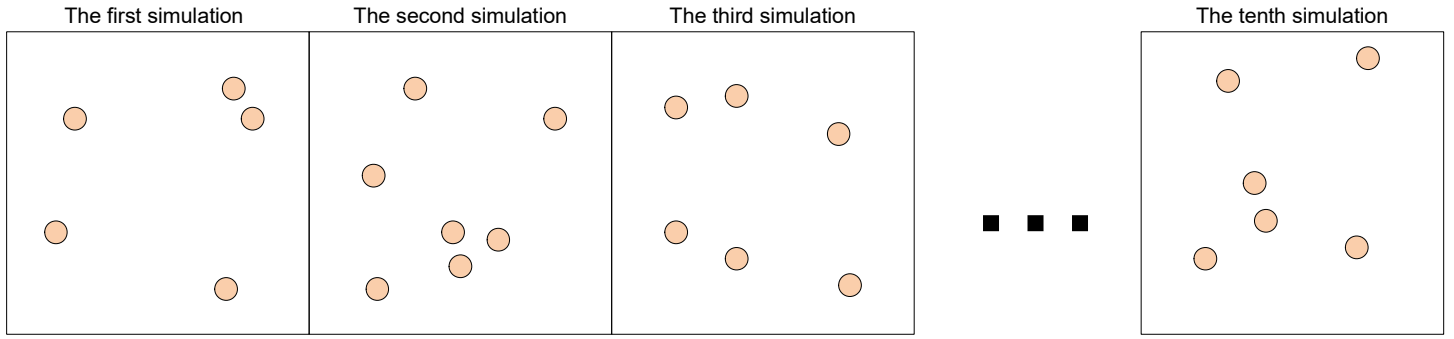




Figure 6.2.3.5 Combine 10 simulation files together by translating 9 of them

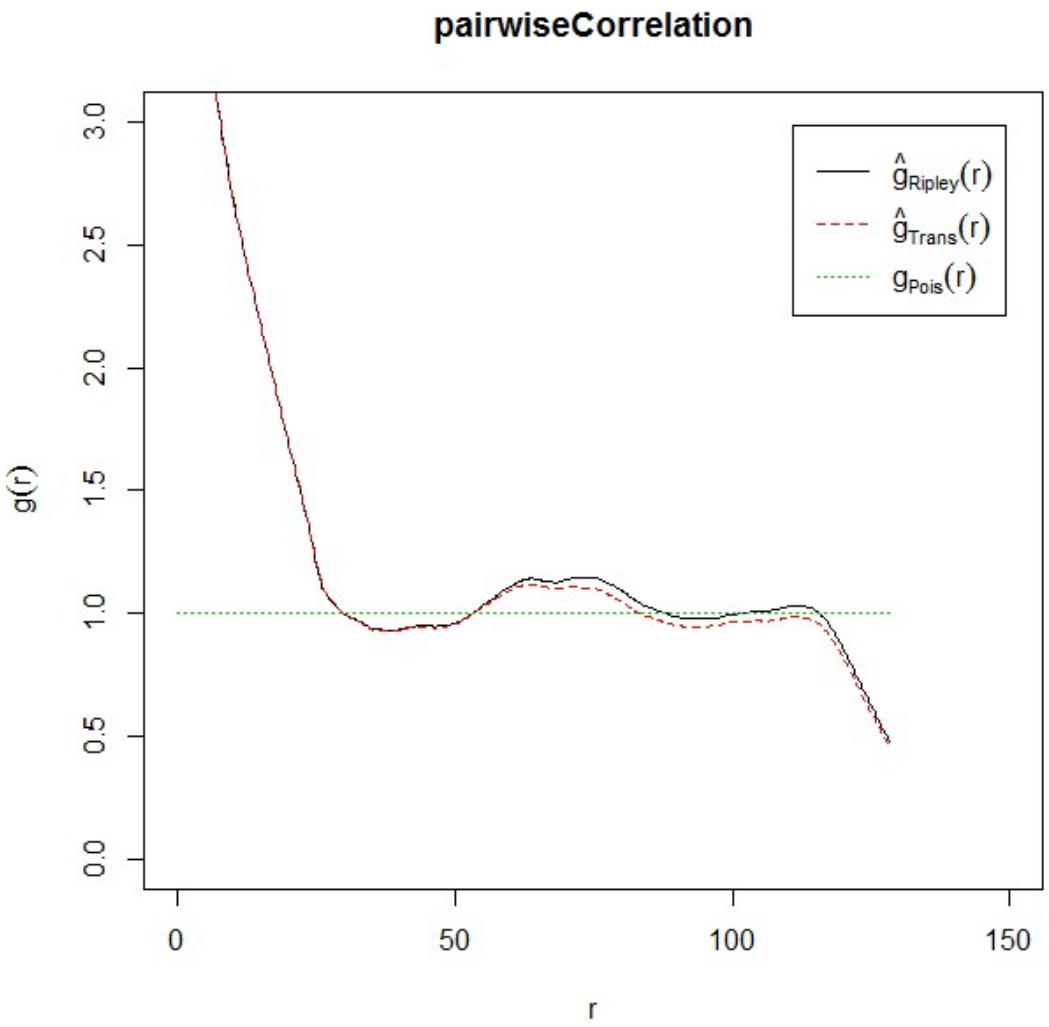



Figure 6.2.3.6 The pairwise correlation curve of combination of 10 simulation output files

The peak near the origin remains in the pairwise correlation function remains, which indicates that the low total number of agent is not the reason. Thus the second assumption is made, that the peak near the origin is caused by a low density. Compared to the number density and pairwise correlation curve of control group, I can see that number density around 0.00012 can produce a high peak near the origin in R.

| | Number density |
|---|---|

| | |
|---|---|
| Start point of control group experiment | 0.00012 |
| Middle point of control group experiment | 0.00014 |
| End point of control group experiment | 0.00013 |

Table 6.2.3.3 number density of control group experiment

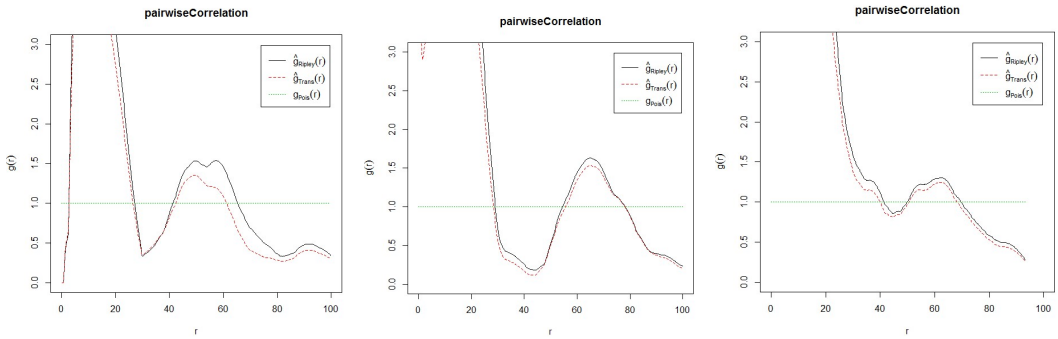



Figure 6.2.3.4 Sample pairwise correlation function of the start (left), middle (middle) and end (right) point of control group growth experiment

## 6.2.4 Comparison against the combined 5FU and hypoxia experiment

Following the approach of Chapter 5, and in spite of the difficulties faced in calibrating the model with experimental data of hypoxia group, I test the derived parameter setting for each of the 5FU and hypoxia simulations by using the average value for parameters of each simulation group is used for the simulation of combination of 5FU and hypoxia.

The time-lapse image and pairwise correlation of starting, middle and end point of experiment with combination of 5FU and hypoxia is shown in Figure 6.2.4.1.

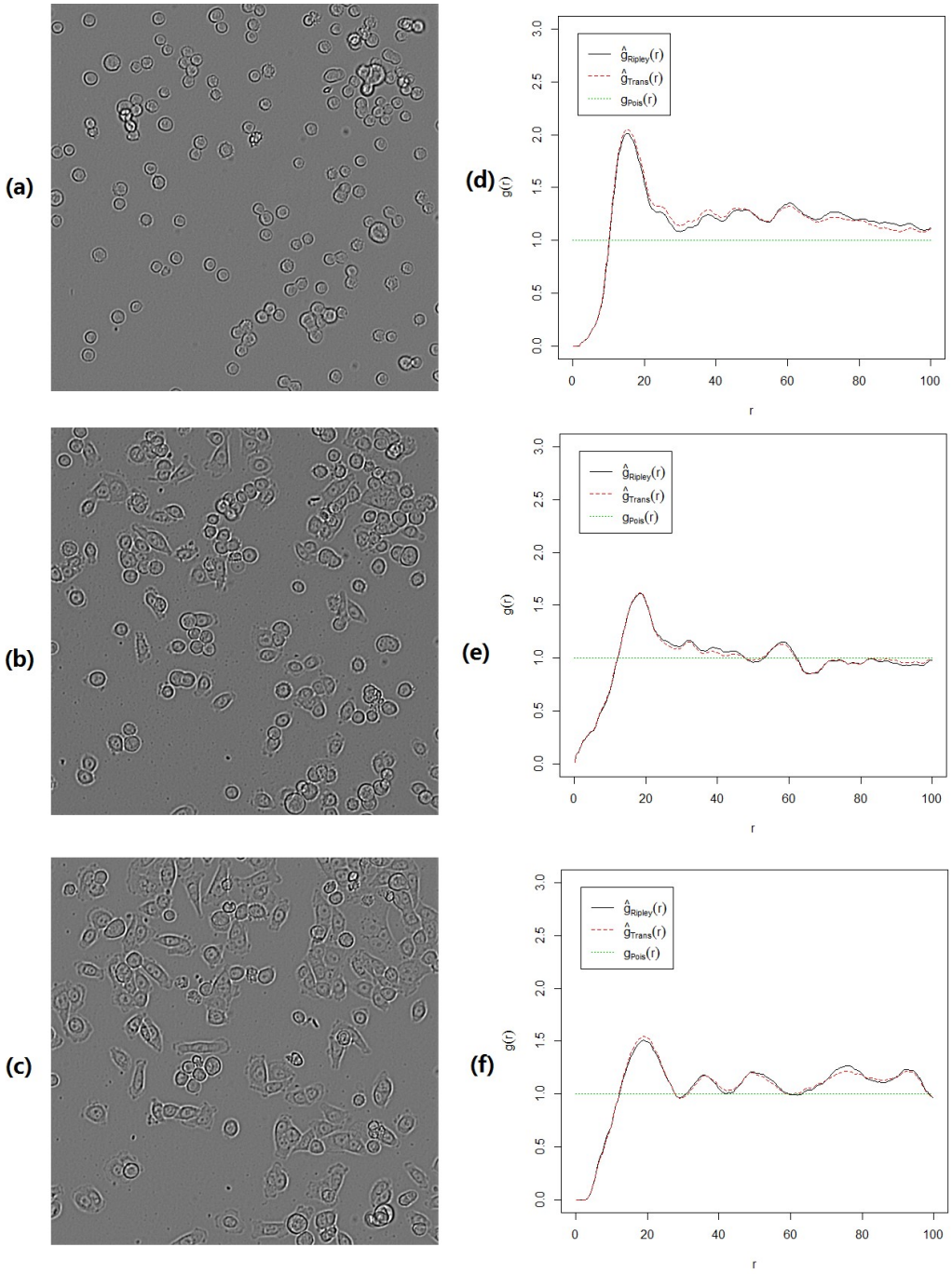



Figure 6.2.4.1 Time-lapse image and pairwise correlation curve of 5FU+hypoxia group of experiment. (a) is the time-lapse image of starting point of experiment, (d) is pairwise correlation function of (a); (b) is the time-lapse image of middle point of experiment, (e) is pairwise correlation function of (b); (c) is the time-lapse image of the ending point of experiment, (f) is pairwise correlation function of (c).

The Figure 6.2.4.1 (a) is the time-lapse image of starting point of experiment, and 6.2.4.1 (d) is the pairwise correlation function of Figure 6.2.4.1 (a). Similar with the end point of 5FU group, as shown in Figure 6.2.4.2, the pairwise correlation function in 6.2.4.1 (d) indicates

one single cluster with all the cells. In this case a circular observation window is needed to correctly display spatial feature of Figure 6.2.4.1 (a). The pairwise correlation function with circular observation window is shown in Figure 6.2.4.3.

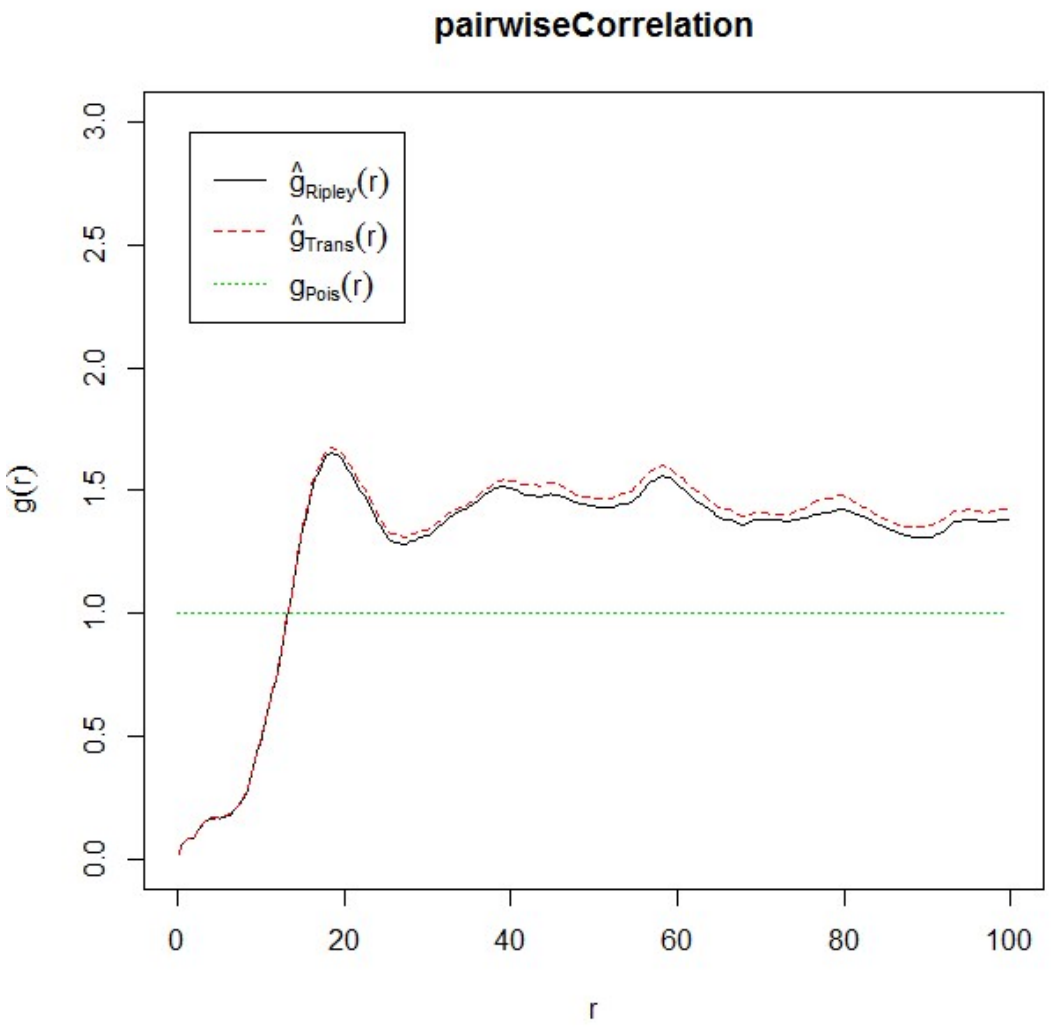



Figure 6.2.4.2 Pairwise correlation function of the end point of experiment with effect of 5FU. From distance=20 to distance=100, the value of $g(r)$ is alway larger than 1, which means all the cells form one large cluster. This means a circular observation window is needed.

The pairwise correlation function with circular observation window, as shown in Figure 6.2.4.3, indicates a peak near distance=20 and a valley near distance=30. The peak also exists in the end point of the combination group on similar distance, with value of $g(r)$ slightly reduced, which means less cells are clustered. There is also a valley in the end point at similar distance.

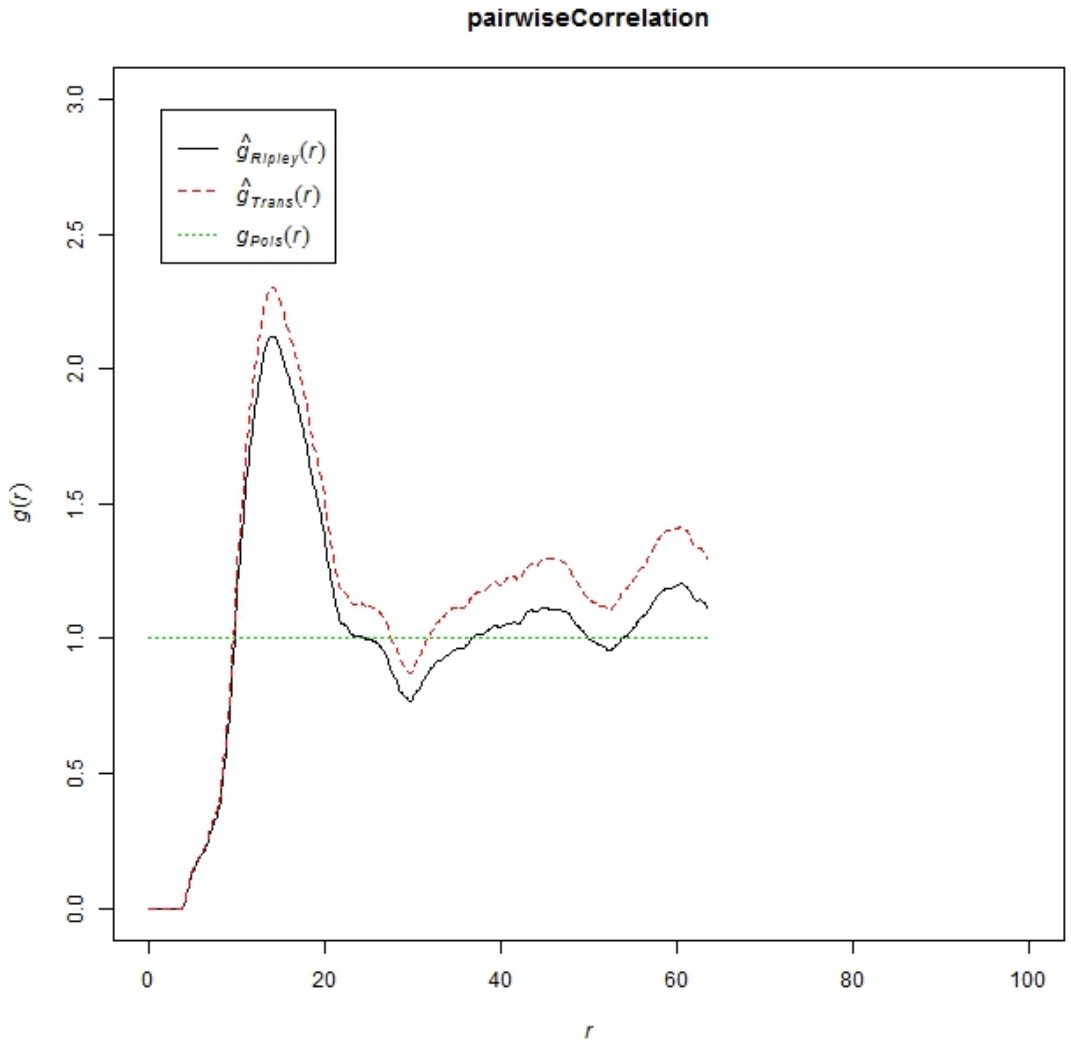



Figure 6.2.4.3 Pairwise correlation function of the start point of experiment with combination of 5FU and hypoxia. With a circular observation window, the range of distance is limited, however the spatial feature within the range is displayed with peak and valley.

Figure 6.2.4.4 shows the starting, middle and end point of simulation under effect of 5FU and hypoxia. Same with the simulation of hypoxia group, there are two types of agents in the simulation of combination group, which are represented by red and green dots. Comparing to the simulation of hypoxia group, the 5FU+hypoxia group contains less agents in the middle (Figure 6.2.4.4 (b)) and end point (Figure 6.2.4.4 (c)).

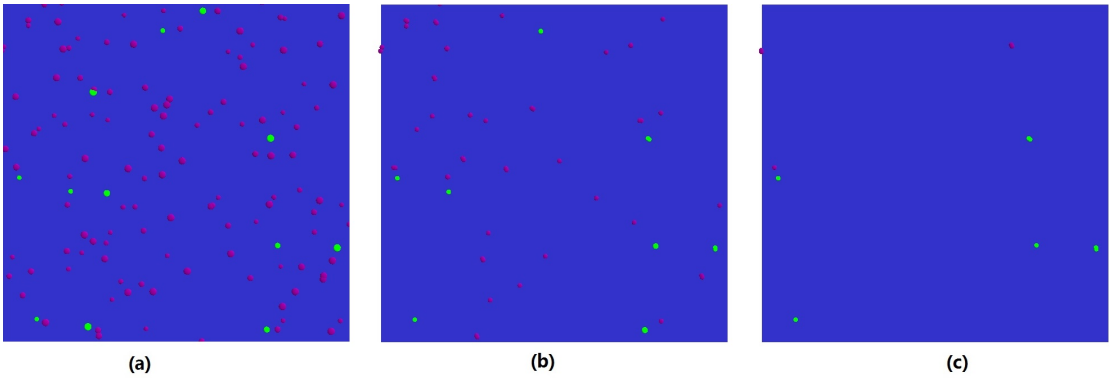



Figure 6.2.4.4 Simulation result of agents under effect of 5FU+hypoxia. The red and green part are two types of agents, while the dark blue part is background. (a) is the start point of simulation, (b) is the middle point of simulation, and (c) is the end point of simulation.

The combination simulation is also run for 10 times, and the output position of agents is also filtered according to the distance. For completeness here I show the average pairwise correlation function for each of 5FU, hypoxia and the 5FU-hypoxia combination simulations is shown in Figure 6.2.4.1, Figure 6.2.4.2 and Figure 6.2.4.3.

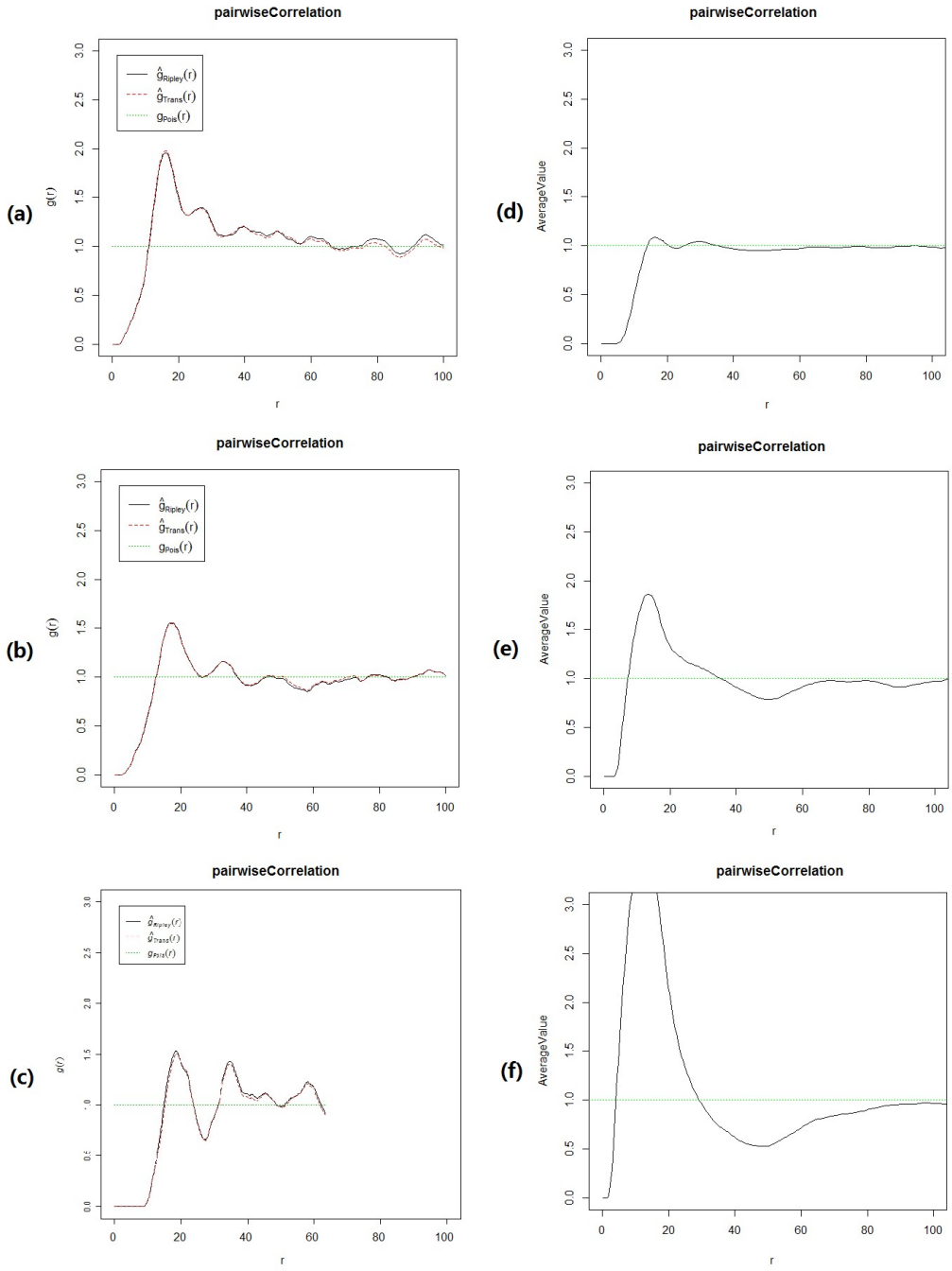



Figure 6.2.4.5 Pairwise correlation function of experimental data (a) - (c), and average pairwise correlation function of 10 simulations (d) - (f) of 5FU group. (a) and (d) are the starting point of experimental and simulation data; (b) and (e) are middle point of experimental and simulation data; and (c) and (f) are end point of experimental and simulation data.

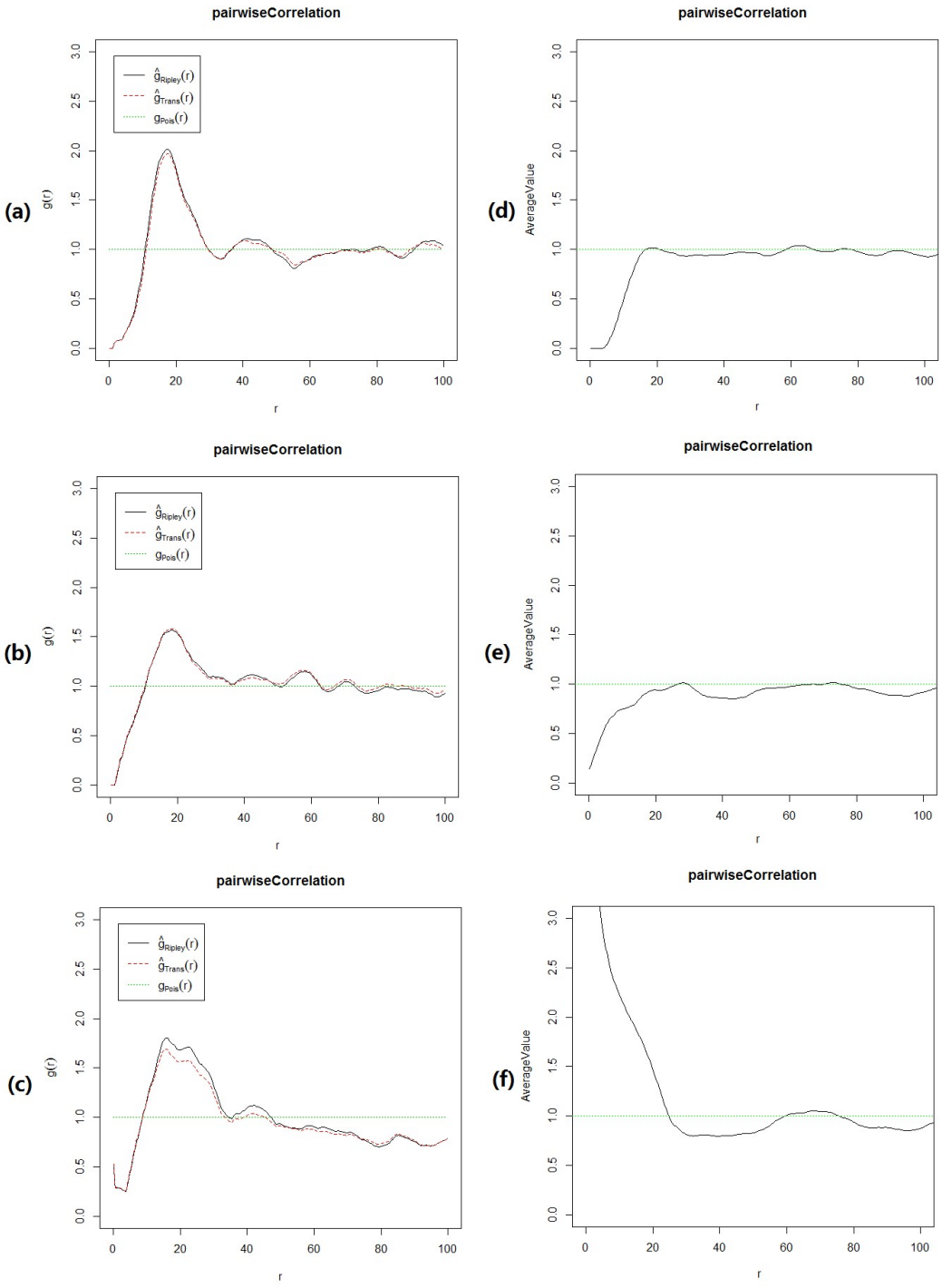



Figure 6.2.4.6 Pairwise correlation function of the starting, middle and end point of experimental data (a) - (c) and average of 10 simulations of hypoxia (d) - (f).

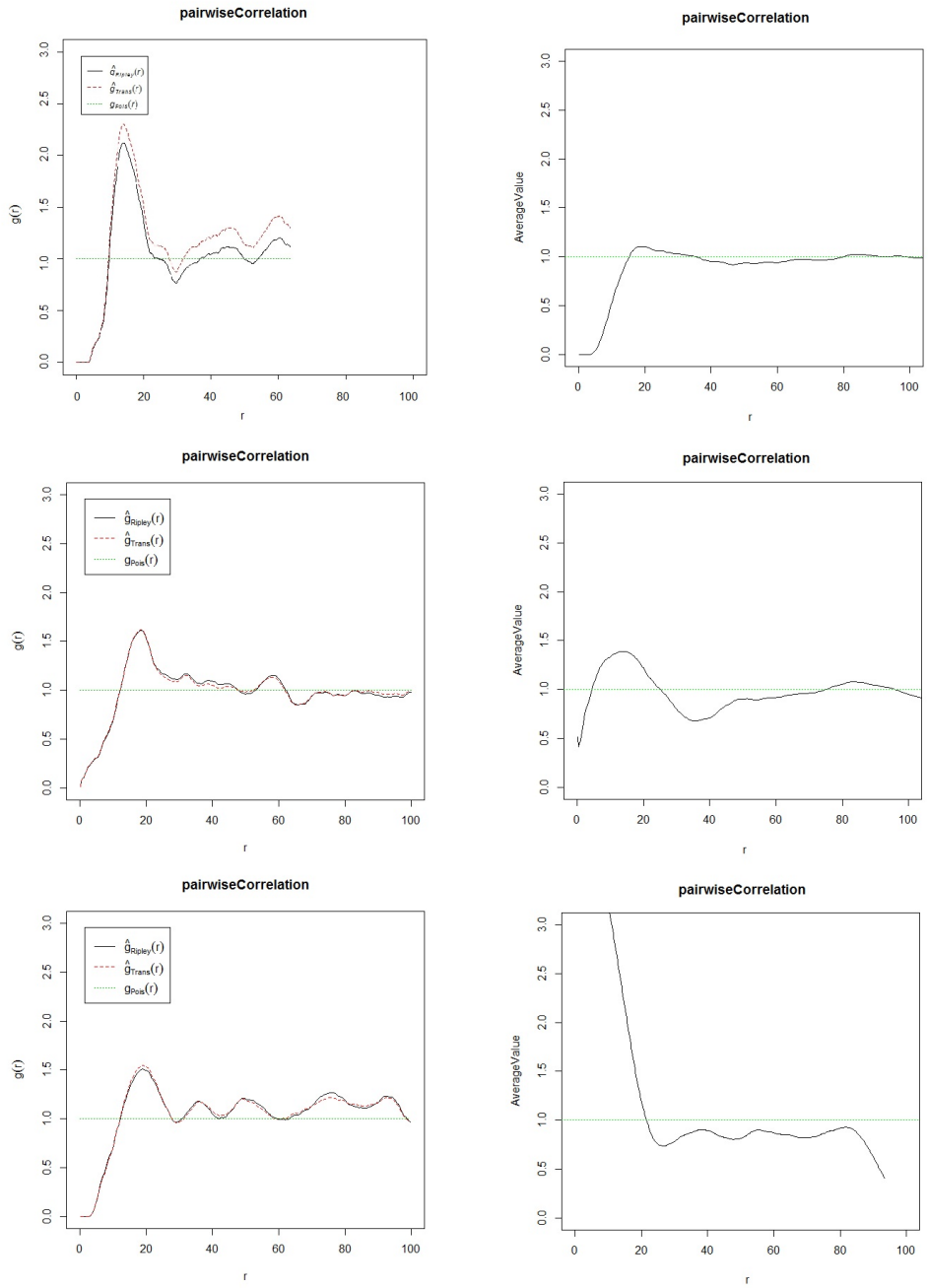



Figure 6.2.4.7 Pairwise correlation function of experimental data (a) - (c), and average pairwise correlation function of 10 simulations (d) - (f) of combination group (5FU+hypoxia). (a) and (d) are the starting point of experimental and simulation data; (b) and (e) are middle point of experimental and simulation data; and (c) and (f) are end point of experimental and simulation data.

From Figure 6.2.4.7, it can be observed that, the simulation starts from random distribution, while the experimental data starts from a clustered distribution. Although the number density at the end point of simulation is to low to give out useful information, the simulation is able to reproduce the spatial feature of a peak near

distance=20 and a valley near distance = 30 in the middle point.

In all three starting points, the simulation groups start from a random distribution. In the middle point of simulations, the peak of the combination simulation appears in a similar position to the 5FU simulation although with a lower peak value, and can be considered as average value of 5FU and hypoxia simulations. At the end point, the combination simulation contains fewer agents than the hypoxia simulation, and the peak near the origin appears. In the end point of all three simulations, the curves are under 1.0 on distance>20. As shown in initial test simulations discussed in section 6.2.1, in the combination simulation the agents spread over a shorter range because of low number density. From the comparison of three simulation and experiment groups, it is shown that the simulation result of 5-FU group produces the most similar pairwise correlation curve with experiment; the hypoxia simulation produces the least similar curve; and the combination group shows a result that is more similar than hypoxia group but less similar than 5-FU group.

# Conclusion

In chapter 6, I first explored the features of experimental output (time-lapse image data) and chose the pairwise correlation curve as the tool to represent the distribution of cells. Considering the experimental output as the result of cell interaction, I then run simulations with various parameter settings. The resulting pairwise correlation curves from simulations are presented and compared with the corresponding experimental curves. As expected, by changing simulation parameters, pairwise correlation curves with different shapes are produced. Also as expected, the range that

adhesion force takes effect does affect the result pattern. I did not expect the speed of agents to be a factor in agent distribution, however after discuss with biological experts I reduced the velocity of agents as effect of drug in simulation and had a good fit with experimental result (pairwise correlation curve). Finally, the balancing point (where the contact force has same value and opposite direction of adhesion force) does not make much difference to the pairwise correlation curve. The possible reason is that the balancing point works with agents of exactly the same shape, and with agents of various cell cycles the size of each agent varies, which introduces a disturbance to the distance on which agents balance.

Besides, only when I reached the simulation phase did it become evident that the uniform density in the time-lapse images of 5FU, hypoxia and combination group in experiment is lower than expected from the population growth experiments. If I are able to repeat the growth experiment under improved conditions I may be able to explain this.

# Chapter 7 Conclusion and Future work

## 7.1 Research Hypotheses

The aim of this thesis was to develop an individual-based model of populations cells and inter-cell interactions. The model focused on characterising the physical properties of cells and how these drive inter-cell interaction dynamics. The biological behaviour of cells was captured in terms of changes to the physical property of cells, and assumes that the physical model can be used to predict biological phenomenon. This assumption led to the following hypotheses:

1 An *in silico* physical model can reproduce important behaviours of in experimental systems of human cells;

2 The model can be parameterised with biologically interpretable values to model differences arising from experimental conditions for the same biological system;

3 The characterisation of cell distributions by spatial statistics enables prediction of biological phenomena using the physical model.

To test these hypotheses, a physics-based model was developed, as detailed in Chapter 3, underpinned by a set of assumptions emerging from the review of literature in Chapter 2.

The first biological system to be studied was the early stage formation of vascular structure by endothelial cells. This system is tested by *in silico* simulations in Chapter 4, in which ellipsoidal agents representing cells were placed at random positions with random directions on a substrate. In this model, endothelial cells do not divide or die in this process. Note that although no tubular structure is formed the connected agents are considered a vascular structure.

The simulation result is considered as a prediction and is compared with *in silico* simulation of another model. With spatial statistics analysis, the typical distance between pair of agents is considered the typical size of holes in net-shaped structure. A key observation was that agents could indeed self-organise into vascular-like structures simply by manipulating both cell density and physical parameters. This work was on track to explore and ultimately test Hypotheses 1 and 2, with some early confirmation of Hypothesis 1 established. However, due to lack of data streams for *in vitro* experiments of early stage of vascular formation, as noted in Chapter 1, the physical model was repurposed to describe the interactions among tumour cells in Chapter 4.

Through a structured analysis of the physical properties of cells in both vascular formation and cancer cell interactions, I assumed that both biological systems contained similar cell-cell physical interaction. The vascular formation model was then systematically converted for use with cancer cells in Chapter 4 in order to fully explore Hypotheses 1 to 3.

The major difference between the early-stage study of vascular formation and cancer cell interactions (over a longer time period) is that cells may divide or die in cancer cell population growth experiments. Thus a cell cycle is added to model, in which the parameter to determine if an agent is to divide or die is used to reflect the population growth of cancer cells under different environment condition. The cell cycle model is discussed in Chapter 5 and is based on changes to the physical parameters of the cell. The predictions of model and experiment result were compared by population growth curves for a control condition (Hypothesis 1). The model was then used to predict population growth curves in a range of different experimental conditions (Hypothesis 2).

Two parameters in the model are considered to have effect on pairwise correlation curve, and the discussion is in Section 6.2.1. In consideration of Hypothesis 3, the physical model (with cell cycle part) is used to predict the spatial distribution of cells in *in vitro* experiments under different experimental conditions in Sections 6.2.2, 6.2.3 and 6.2.4.

## **7.2** Conclusion

The Hypothesis 3 is proved in Chapter 6. The characterisation of both cell distribution from experiments and predicted distribution made by the model is demonstrated by pairwise correlation curve. Figure 7.2.1 shows the time-lapse image from experiment and corresponding pairwise correlation curve. I can see that the pairwise correlation curves of 5-FU, hypoxia and combination (experimental) group show both common and specific features. In all three groups, the first peak of pairwise correlation curves appear around distance=20. As I explain the first peak relative with typical size of cell,

and in three experiments the same type of cell is used, this result is reasonable.

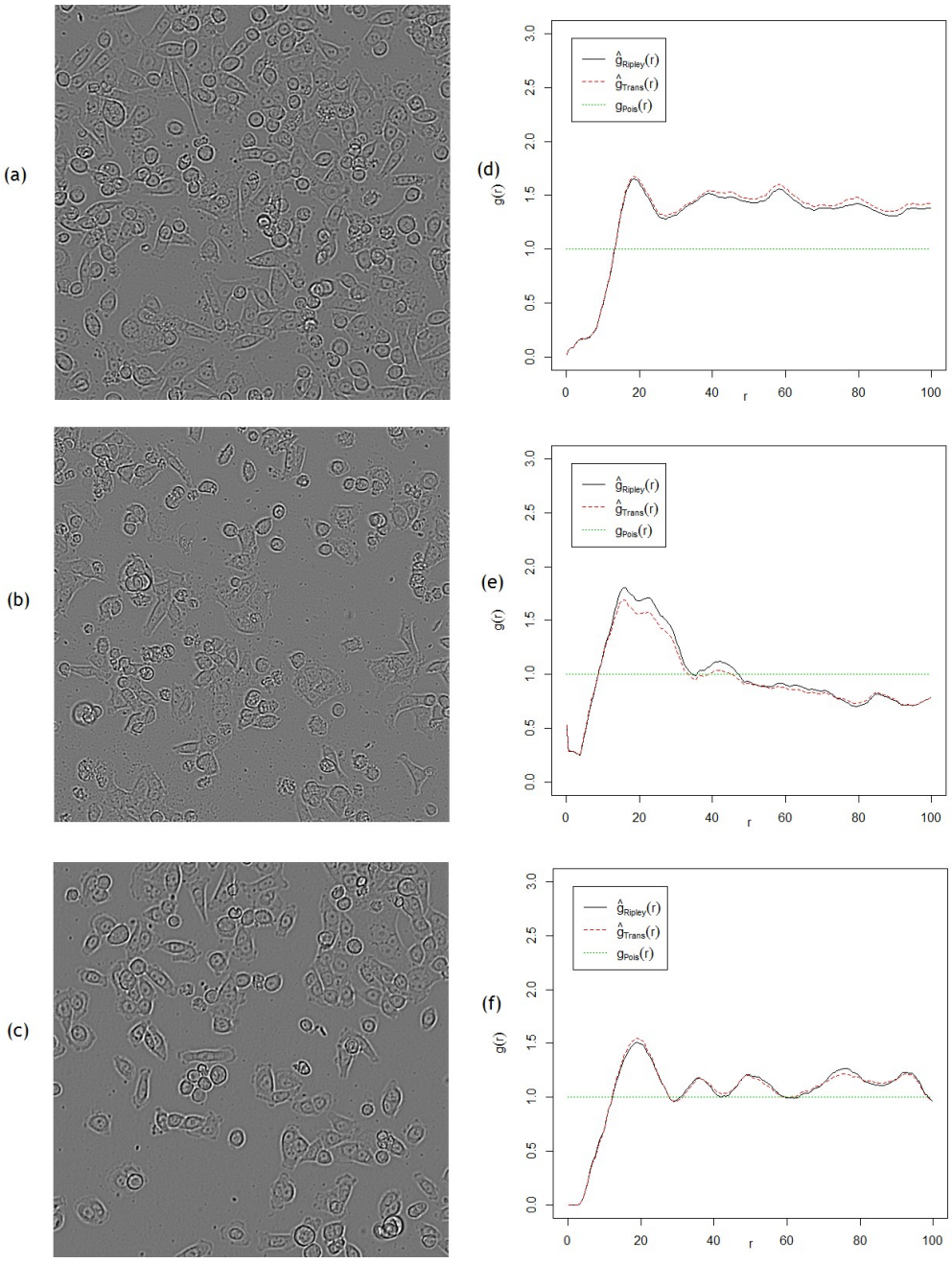



Figure 7.2.1 (a) last image of cancer cell growth experiment with 5-FU; (b) last image of cancer cell growth experiment in hypoxic condition; (c) last image of cancer cell growth experiment with 5-FU in hypoxic condition; (d) is the pairwise correlation curve of (a); (e) is the pairwise correlation curve of (b); (f) is the pairwise correlation curve of (c).

From experimental images, I can see that the 5-FU group shows clusters which are looser, while the hypoxia group contains compact clusters. With data from 5-FU

experimental group, the first peak of pairwise correlation curve has lower value than the hypoxia group. The experimental images also show much closer clusters in 5-FU image, while the hypoxia image contains sparsed clusters. From corresponding pairwise correlation curve I can see that after the first peak the 5-FU curve stays above the line $g = 1$ which indicates random distribution; while the hypoxia curve falls below $g = 1$ after the first peak.

These features of cell distribution are used to calibrate the physical model. When the calibrated model produces predictions, they are also analysed by pairwise correlation function to be compared with experimental data. (H 3: The characterisation of cell distributions by spatial statistics enables prediction of biological phenomena using the physical model)

Hypothesis 2 is proved in Chapters 5 and 6. In Chapter 5, the model is parameterised to predict the number of cells in cancer cell growth experiments. The key parameter in the model is the chance that agents divide to two agents when they finish the cell cycle. The control group provided data to calibrate the model as a baseline. Then the effect of 5-FU and hypoxia on chance of division was modelled and added to physical model separately. Thus the physical model with 5-FU or hypoxia mechanism contained parameter that has biological meaning. From Figure 7.2.2 (a)-(c) I can see that as the baseline, the prediction of control group shows similar trend of increase/decrease of number of cells; the prediction of 5-FU group differs with experimental data of Day1 and Day2 but has better match with data of Day3, 4, and 5; the prediction of hypoxia group differs with data of Day2 but shows good match with experimental data of Day1 and Day3 and similar trend of increase/decrease of Day4

and 5. I also tested the parameters by adding both mechanisms of 5-FU and hypoxia to physical model together and comparing the prediction with corresponding *in vitro* experiment (Figure 7.2.2 (d)), and the prediction shows similar trend of increase/decrease with experimental data. Therefore the model with parameter of different value can predict number of cells of same type under different condition.

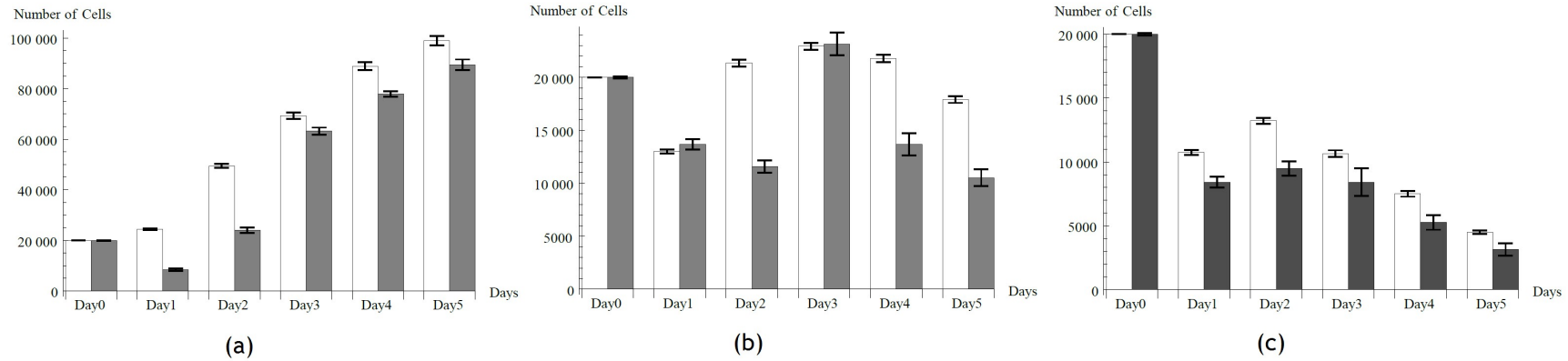



Figure 7.2.2 (a): Experimental and simulation growth curves of 5-FU group; (b): Experimental and simulation growth curves of hypoxia group; (c): Experimental and simulation growth curves of combination of 5-FU and hypoxia group. Greyscale bars are experimental data, and white bars are simulation result.

In Chapter 6, with mechanisms modelled in Chapter 5, the physical model is used to predict the distribution of cancer cells in *in vitro* experiment. The range of adhesion force, the point where adhesion force and contact force balance and the speed that agents move are considered as three key parameters. As the image from control group did not reflect the number of cells in experiment, the 5-FU group is used as the baseline (discussed in Section 6.2.4). By changing values of the three key parameters, the model predicted the distribution of cells as shown in Figure 7.2.3.

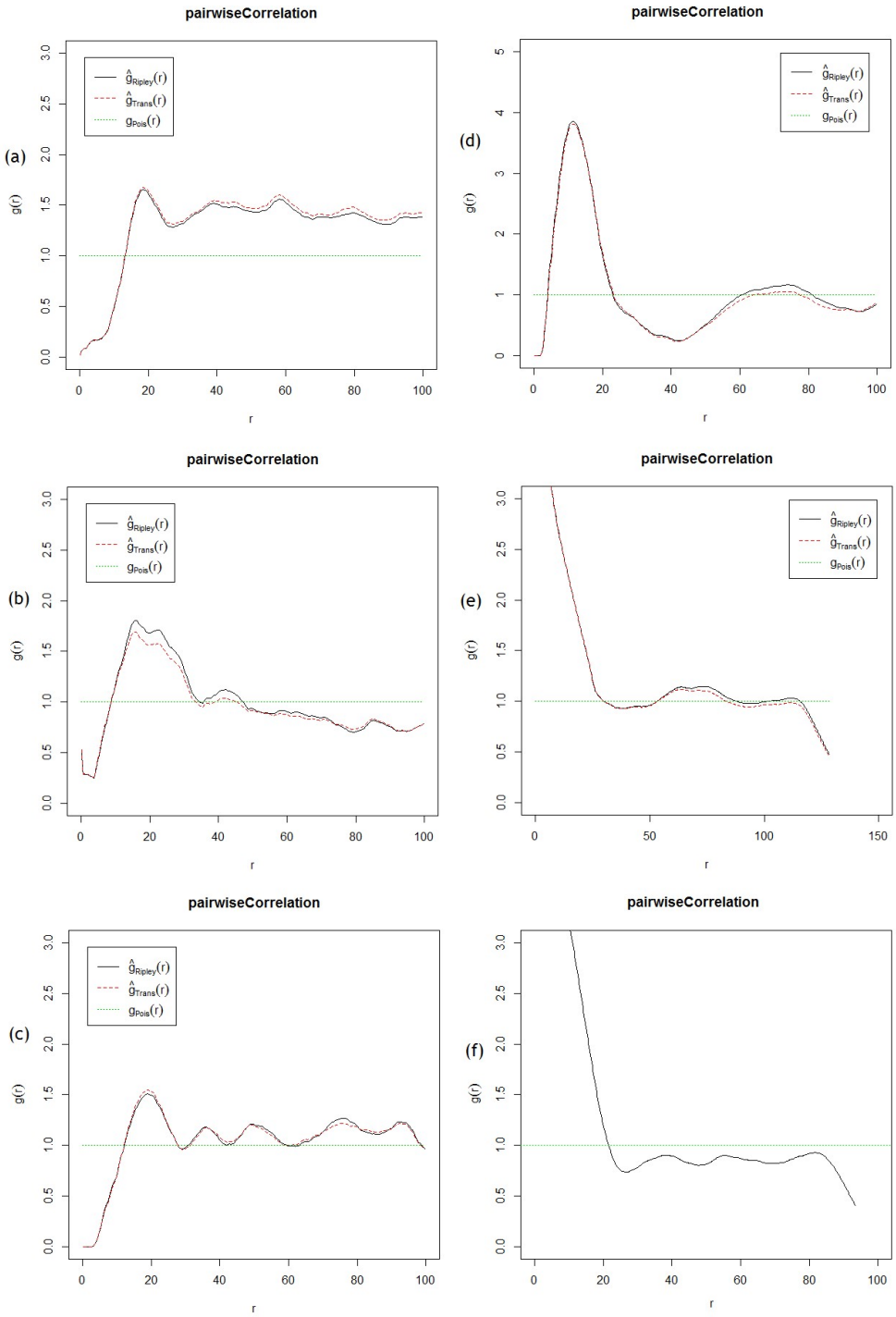



Figure 7.2.3 (a) pairwise correlation curve of last image of cancer cell growth experiment with 5-FU; (b) pairwise correlation curve of last image of cancer cell growth experiment in hypoxic condition; (c) pairwise correlation curve of last image of cancer cell growth experiment with 5-FU in hypoxic condition; (d) pairwise correlation curve of simulation of cancer cell growth with 5-FU (Note that to show actual value of peak, the scale of this figure is different from others); (e) pairwise correlation curve of simulation of cancer cell growth experiment in hypoxic condition; (f) pairwise correlation curve of simulation of cancer cell growth experiment with 5-FU in hypoxic condition.

Comparing the pairwise correlation curves produced by simulation and experimental data, I can see that the first peak in both simulation and experimental curves appears around distance=20. The prediction of 5-FU group shows a higher first peak; although the predicted curve is lower than experimental curve after the first peak, the predicted curves shows increased trend of gathering of agents on long distance (distance>60). The prediction of hypoxia group has a higher first peak and lower value on longer distance, which reflects a different feature of distribution. (H2: The model can be parameterised with biologically interpretable values to model differences arising from experimental conditions for the same biological system)

In Chapter 4, the prediction of vascular formation was also explored. Due to lack of dataflow, the prediction was not calibrated with any *in vitro* experiment data. However the *in silico* simulation still shows the emergent behaviour of randomly placed identical agents. The agents connect with each other alone the longest axis and form net-shaped structure. By changing the number density of agents, the physical model can be used to generate thick but unstable structures or thin but stable structure, as shown in Figure 7.2.4.

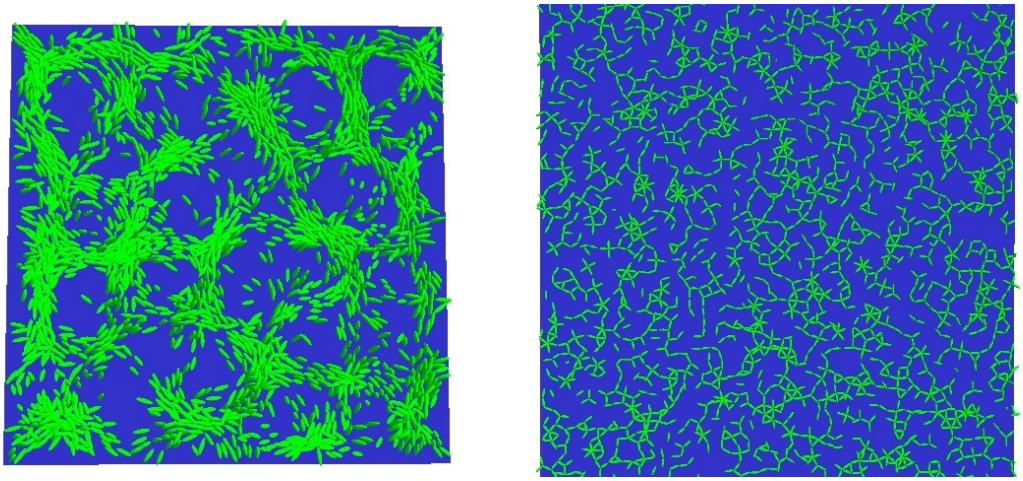


Figure 7.2.4 Thick and thin net-shaped structure

In Chapter 5 the physical model successfully reproduced the growth curve of *in vitro* growth experiment of cancer cells under various conditions. In Chapter 6 the model is used to reproduce the distribution of cells in *in vitro* growth experiment of cancer cells under various conditions. Therefore the Hypothesis 1 is proved by Chapters 4, 5 and 6. (H1: An *in silico* physical model can reproduce important behaviours of in experimental systems of human cells)

## 7.3 Future work

As discussed in Chapter 2, following the process-based modeling approach the mechanism of unknown part of biology phenomenon can be assumed and modeled, and various assumptions can be made for the same phenomenon, such as cancer cell growth in hypoxic environment in Chapter 5. By comparing the in silico result of model, the assumptions can be tested and compared.

This physical model is proved to be able to predict the vessel formation and be reused to predict the spatial distribution of cancer cell in *in vitro* growth experiment. With biological data such as cell size, cell shape, etc. this model is able to predict behaviours of various cell types, which requires physical and biological measurements from experiments. In Section 6.2.3 it was discussed about the problem of cell recognition in CellProfiler. In CellProfiler, the cells within clusters was failed to be recognised, instead, the cluster was identified as a single cell. Therefore errors about both the data about cell positions and number of cells are introduced. In future, with more accurate methods to identify and track cell, such as (Giesen *et al*. 2014), the model can be better calibrated and produce better predictions.

With support of better image process method, the model can also be used to predict more complex phenomenon. A straightforward extension of existing cancer cell growth experiment is that, instead of planting cancer cells on petri-dish to form 2-dementional pattern, the cancer cells can be planted in centre of gel to form 3-dementional pattern (spheroids). There are also problems that may be brought about prediction of spheroidal growth. First, it is hard to directly analysis the pattern formed by spheroidal growth. Thus instead of directly comparing the pattern and prediction, the spheroidal cancer tissue needs to be sliced. Second, in Chapter 5 to model the cancer growth in hypoxic condition, a second cell type is introduced to reflect the effect of hypoxia on both cell growth rate and length of cell cycle. In spheroidal growth the cancer cells in the centre of spheroids are hypoxic, while the cells on surface remain normal. This demands a method in spatial statistic to describe the distribution of mixed type of cells, including the spatial distribution of each type and the spatial relationship between types. The marked point process (Ecke *et al*. Analysis of spatial correlations in marked point processes (ppt)) could be a useful tool because it is to describe the spatial distribution of mixed types of particles. Similar with pairwise correlation function, marked correlation function produces curves that present the spatial distribution in graph.

As shown in literature review (Chapter 2), cancer is a complex system. Apart from predicting two cell types, as an agent-based model, it is relatively easy for this model to predict more complex systems, which can be mixed types of cancer cells, same type of cancer cells in different stage of mutation, or mixed cancer cells and vascular structure. Combining with the advanced cell tracking and identification method, and

proper spatial statistic tools, this model may be used to predict complex, tissue-level phenomenon.

Ultimately, this model will be able to predict the behaviours of mixed type of cancer cells growing in 3-dimension with vascular structure, showing how spheroidal growth of cancer cells produces hypoxic condition in inner layers of cells and how vascular structure can reduce the level of hypoxia.

# Appendix A: Dimension analysis

Dimension analysis requires substitution of the corresponding dimension to the right hand side of all the main equations for forces and torques, after that the value of parameters can be estimated. As the first step, I test the correction of dimension for all equations presenting physical interactions in my model. The estimation of value of parameters is the second step.

I define three basic physical entities: unit length $\Delta x$, unit time $\Delta t$, and unit density $\rho_c$. Each physical entity in my model can be represented by the dimension version of itself, which is in form of the physical entity divided by its dimension. Therefore there are two methods to find the dimension version of a physical entity. One is from the physical definition of this physical entity. For example, the contact force, as a force it can be defined by Newton's law $\mathbf{F} = m\mathbf{a}$, in which the dimension of mass $m$ is volume (cubical of length) multiplied by density, or $\Delta x^3 \rho_c$; the dimension of acceleration $\mathbf{a}$ is the change of velocity in unit time, or $\frac{\Delta x}{\Delta t^2}$. Thus from the definition of force, the dimension of contact force is $\Delta x^3 \rho_c \cdot \frac{\Delta x}{\Delta t^2}$, or $\frac{\Delta x^4 \rho_c}{\Delta t^2}$.

The other method to find dimension of a physical entity is by deriving the equation from which this physical entity is calculated. Again take the contact force as an example, from its formula $\mathbf{F}^{con} = K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}} + \left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2 \left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}} \cdot \left(1-\sqrt{\mathbf{\Phi}}\right)^{\frac{3}{2}} \cdot \mathbf{\Phi}^{-\frac{1}{2}} \cdot \left(\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right)$, the dimension of contact force can be found by writing all the entities in right hand side of equation with their dimension version. If from two methods the same dimension of entity is derived, then I consider this physical entity has correct dimension.

## Step 1: Dimension test for all physical entities in the model

In the following section, I derive the dimension for all the physical entities in my model. Although some steps look simple or meaningless, this step is to make sure the equations are not malformed during complex equation derivation in Chapter 3 to let a force dimensionally equal to a distance. That is why this step is important and necessary.

Firstly, the displacement between two ellipsoids **A** and **B** is $\mathbf{s} = \left(\mathbf{r}_A - \mathbf{r}_B\right)$ (defined in equation (44)), I can let $\Delta x$ be unit length, so that s turns to

$$\tilde{\mathbf{s}} = \frac{\mathbf{s}}{\Delta x} \tag{1}$$

In which $\tilde{\mathbf{s}}$ is the dimensionless version of **s**. Note that $\Delta x$ is a scalar quantity and **s** is a vector, and so after the division, $\tilde{\mathbf{s}}$ is also a vector. In the following part of this section, I will use $\tilde{P}$ to represent the dimensional version for any physical quantity $P$.

## Contact potential:

From its definition, it is known that the contact potential is a scalar quantity. I calculate it here step by step. From equation (46) it is known that $\Phi$ is proportional to $\lambda_0$, displacement **s** and matrix **G**. As $\lambda_0$ is a real value between zero and one, it is scalar as well. The definition of **G** in equation (20)**Error! Reference source not found.** shows it is proportional to matrix $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$. Because of equation (11) and (12)

$$\mathbf{A}^{-1} = a_1^2\mathbf{u}_1 \otimes \mathbf{u}_1 + a_2^2\mathbf{u}_2 \otimes \mathbf{u}_2 + a_3^2\mathbf{u}_3 \otimes \mathbf{u}_3 \tag{2}$$

$$\mathbf{B}^{-1} = b_1^2\mathbf{v}_1 \otimes \mathbf{v}_1 + b_2^2\mathbf{v}_2 \otimes \mathbf{v}_2 + b_3^2\mathbf{v}_3 \otimes \mathbf{v}_3 \tag{3}$$

$a_1$, $a_2$, $a_3$ and $b_1$, $b_2$, $b_3$ are lengths of the ellipsoid semi-radii, so their dimension is $\Delta x$, and **u** and **v** are unit matrices, so the dimension of $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ is $\Delta x^2$. Because of the definition of $\mathbf{G}(\lambda)$, the dimension $\mathbf{G}(\lambda)$ is $\Delta x^2$ as well. Therefore the dimension of $\mathbf{G}^{-1}(\lambda_0)$ is $\frac{1}{\Delta x^2}$. So

$$\tilde{\boldsymbol{\Phi}}(\mathbf{A},\mathbf{B},\mathbf{s}) = \Delta x \cdot \frac{1}{(\Delta x)^2} \cdot \Delta x \cdot \boldsymbol{\Phi}(\mathbf{A},\mathbf{B},\mathbf{s}) = \boldsymbol{\Phi}(\mathbf{A},\mathbf{B},\mathbf{s}) \tag{4}$$

Thus, the dimension of $\boldsymbol{\Phi}(\mathbf{A},\mathbf{B},\mathbf{s})$ is 1, or I can say that $\boldsymbol{\Phi}(\mathbf{A},\mathbf{B},\mathbf{s})$ is non-dimensional, i.e. it is scalar.

## Contact force:

The next object to test is the contact force. Consider the definition of force in general, $\mathbf{F} = m\mathbf{a}$, so the dimension of force should be a multiple of mass and acceleration. Then mass can be considered as the multiple of its volume and its density. It is easy to determine the dimension of volume is $\Delta x^3$. I let $\rho_c$ be the unit density, so that the dimension of mass is $\rho_c \Delta x^3$. The acceleration, on the other hand, equals distance divided by square of time. I let $\Delta t$ be the unit time, then the dimension of acceleration of **a** is $\frac{\Delta x}{\Delta t^2}$. And the dimension of force should be $\rho_c \Delta x^3 \cdot \frac{\Delta x}{\Delta t^2} = \rho_c \frac{\Delta x^4}{\Delta t^2}$.

From equation (101),

$$\mathbf{F}^{con}=K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}}+\left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2\left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}\cdot\left(1-\sqrt{\mathbf{\Phi}}\right)^{\frac{3}{2}}\cdot\mathbf{\Phi}^{-\frac{1}{2}}\cdot\left(\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right) \tag{5}$$

It is known that $\tilde{a}_i=\frac{a_i}{\Delta x}\left(i=1,2,3\right)$, $\tilde{b}_i=\frac{b_i}{\Delta x}\left(i=1,2,3\right)$, $\tilde{\mathbf{\Phi}}=\mathbf{\Phi}$, $\tilde{\mathbf{G}}=\frac{\mathbf{G}}{\Delta x^2}$, so that $\tilde{\mathbf{G}}^{-1}=\frac{\mathbf{G}^{-1}}{\frac{1}{\Delta x^2}}$.

The elastic parameter $K$ is defined in equation (74) as

$$\begin{cases}K=\dfrac{4}{3\pi\left(k_1+k_2\right)}\\ k_1=\dfrac{1-v^2}{\pi E_1}\\ k_2=\dfrac{1-v^2}{\pi E_2}\end{cases} \tag{6}$$

From their definition, dimension of the (name of physical quantity) $E_1,E_2$ is $\frac{\rho_c\Delta x^2}{\Delta t^2}$, dimension of $v$ is 1, i.e.

$$\tilde{E}=\frac{E}{\frac{\rho_c\Delta x^2}{\Delta t^2}} \tag{7}$$

Thus the dimension of $K$ is $\frac{\rho_c\Delta x^2}{\Delta t^2}$. Plus $\tilde{\lambda}_0=\frac{\lambda_0}{1}$. Therefore

$$\tilde{\mathbf{F}}^{con}=\tilde{K}\left[\left(\tilde{a}_1\tilde{a}_2\tilde{a}_3\right)^{\frac{1}{3}}+\left(\tilde{b}_1\tilde{b}_2\tilde{b}_3\right)^{\frac{1}{3}}\right]^2\cdot\left(\tilde{a}_1\tilde{a}_2\tilde{a}_3\tilde{b}_1\tilde{b}_2\tilde{b}_3\right)^{\frac{1}{6}}$$
$$\cdot\left(1-\sqrt{\tilde{\Phi}}\right)^{\frac{3}{2}}\cdot\tilde{\Phi}^{-\frac{1}{2}}\cdot\left(\tilde{\lambda}_0\left(1-\tilde{\lambda}_0\right)\tilde{\mathbf{G}}^{-1}\left(\tilde{\lambda}_0\right)\tilde{\mathbf{s}}\right)$$
$$=\frac{K}{\frac{\rho_c\Delta x^2}{\Delta t^2}}\left[\left(\frac{a_1}{\Delta x}\cdot\frac{a_2}{\Delta x}\cdot\frac{a_3}{\Delta x}\right)^{\frac{1}{3}}+\left(\frac{b_1}{\Delta x}\cdot\frac{b_2}{\Delta x}\cdot\frac{b_3}{\Delta x}\right)^{\frac{1}{3}}\right]^2\left(\frac{a_1}{\Delta x}\cdot\frac{a_2}{\Delta x}\cdot\frac{a_3}{\Delta x}\cdot\frac{b_1}{\Delta x}\cdot\frac{b_2}{\Delta x}\cdot\frac{b_3}{\Delta x}\right)^{\frac{1}{6}}$$
$$\cdot\left(1-\sqrt{\frac{\mathbf{\Phi}}{1}}\right)^{\frac{3}{2}}\cdot\left(\frac{\mathbf{\Phi}}{1}\right)^{-\frac{1}{2}}\cdot\left(\lambda_0\left(1-\lambda_0\right)\frac{\mathbf{G}^{-1}}{\frac{1}{\Delta x^2}}\left(\lambda_0\right)\frac{\mathbf{s}}{\Delta x}\right)$$
$$=\frac{1}{\frac{\rho_c\Delta x^4}{\Delta t^2}}\cdot K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}}+\left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2\left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}\cdot\left(1-\sqrt{\mathbf{\Phi}}\right)^{\frac{3}{2}}\cdot\mathbf{\Phi}^{-\frac{1}{2}}\cdot\left(\lambda_0\left(1-\lambda_0\right)\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right)$$
$$=\frac{\mathbf{F}^{con}}{\frac{\rho_c\Delta x^4}{\Delta t^2}} \tag{8}$$

From equation (203), the equation for contact force has no dimension problem.

## Contact torque:

From definition of toque, $\boldsymbol{\tau}^{Con}=\mathbf{F}\times\mathbf{L}$, in which $\mathbf{L}$ is the force arm. The dimension of torque is $\frac{\rho_c\Delta x^5}{\Delta t^2}$. From equation (134),

$$\boldsymbol{\tau}^{Con}=-\lambda_0\left(1-\lambda_0\right)^2K\left[\left(a_1a_2a_3\right)^{\frac{1}{3}}+\left(b_1b_2b_3\right)^{\frac{1}{3}}\right]^2\left(a_1a_2a_3b_1b_2b_3\right)^{\frac{1}{6}}\cdot\left(1-\sqrt{\Phi}\right)^{\frac{3}{2}}\cdot\Phi^{-\frac{1}{2}}$$
$$\cdot\left[\mathbf{A}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right] \tag{9}$$

I already know that $\tilde{\lambda}_0=\frac{\lambda_0}{1}$, $\tilde{K}=\frac{K}{\frac{\rho_c\Delta x^2}{\Delta t^2}}$, $\tilde{a}_i=\frac{a_i}{\Delta x}\left(i=1,2,3\right)$, $\tilde{b}_i=\frac{b_i}{\Delta x}\left(i=1,2,3\right)$,

$\tilde{\mathbf{\Phi}}=\mathbf{\Phi}$, $\tilde{\mathbf{A}}^{-1}=\frac{\mathbf{A}^{-1}}{\Delta x^2}$, $\tilde{\mathbf{G}}=\frac{\mathbf{G}}{\Delta x^2}$, so that $\tilde{\mathbf{G}}^{-1}=\frac{\mathbf{G}^{-1}}{\frac{1}{\Delta x^2}}$.

So

$$\tilde{\boldsymbol{\tau}}^{Con}=\frac{\boldsymbol{\tau}^{Con}}{\frac{\rho_c\Delta x^5}{\Delta t^2}} \tag{10}$$

From equation (205), the equation for contact torque has no dimension problem.

## Adhesion energy:

Due to the definition of energy in general, $W = \mathbf{F} \cdot \mathbf{s}$. Thus the dimension of energy is $\frac{\rho_c \Delta x^5}{\Delta t^2}$. I can say the dimension of the adhesion energy is $\frac{\rho_c \Delta x^5}{\Delta t^2}$ as well. let dimension of $\varepsilon$ be $\frac{\rho_c \Delta x^5}{\Delta t^2}$, or I can write it in form of

$$\tilde{\varepsilon} = \frac{\varepsilon}{\frac{\rho_c \Delta x^5}{\Delta t^2}} \tag{11}$$

In which $\varepsilon$ is dimensional quantity, and $\tilde{\varepsilon}$ is non-dimensional quantity. Plus the dimension of $\mathbf{\Phi}(\mathbf{A}, \mathbf{B}, \mathbf{s})$ is 1, then the dimensionless version of equation can be written as

$$\begin{aligned} \tilde{W}^{Ad} &= \tilde{\varepsilon}\sqrt{\tilde{\mathbf{\Phi}}(A, B, s)} \\ &= \frac{\varepsilon}{\frac{\rho_c \Delta x^5}{\Delta t^2}} \cdot \frac{\sqrt{\mathbf{\Phi}(A, B, s)}}{1} \\ &= \frac{\varepsilon\sqrt{\mathbf{\Phi}(A, B, s)}}{\frac{\rho_c \Delta x^5}{\Delta t^2}} \\ &= \frac{W}{\frac{\rho_c \Delta x^5}{\Delta t^2}} \end{aligned} \tag{12}$$

Therefore the equation for adhesion energy has no dimension problem.

## Adhesion force:

The dimension of force is $\frac{\rho_c \Delta x^4}{\Delta t^2}$. The definition of adhesion force is as follows

$$\mathbf{F}^{Ad} = -\varepsilon\lambda_0 (1-\lambda_0) \Phi^{-\frac{1}{2}} \mathbf{G}^{-1}(\lambda_0)\mathbf{s} \tag{13}$$

It is known that $\tilde{\varepsilon} = \frac{\varepsilon}{\frac{\rho_c \Delta x^5}{\Delta t^2}}$, $\tilde{\lambda}_0 = \frac{\lambda_0}{1}$, $\tilde{\mathbf{\Phi}} = \mathbf{\Phi}$, $\tilde{\mathbf{G}}^{-1} = \frac{\mathbf{G}^{-1}}{\frac{1}{\Delta x^2}}$, $\tilde{\mathbf{s}} = \frac{\mathbf{s}}{\Delta x}$

So

$$\begin{aligned} \tilde{\mathbf{F}}^{Ad} &= -\frac{\varepsilon}{\frac{\rho_c \Delta x^5}{\Delta t^2}} \cdot \lambda_0 (1-\lambda_0) \cdot \mathbf{\Phi}^{-\frac{1}{2}} \cdot \frac{\mathbf{G}^{-1}}{\frac{1}{\Delta x^2}} \cdot \frac{\mathbf{s}}{\Delta x} \\ &= \frac{\mathbf{F}^{Ad}}{\frac{\rho_c \Delta x^4}{\Delta t^2}} \end{aligned} \tag{14}$$

There is no dimension problem with adhesion force.

## Adhesion torque:

The dimension of torque is $\frac{\rho_c \Delta x^5}{\Delta t^2}$. Equation for adhesion torque is as follows,

$$\boldsymbol{\tau}^{Ad} = \varepsilon\lambda_0\left(1-\lambda_0\right)^2 \Phi^{-\frac{1}{2}}\left[\mathbf{A}^{-1}\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right]\times\left[\mathbf{G}^{-1}\left(\lambda_0\right)\mathbf{s}\right] \tag{15}$$

It is known that $\tilde{\varepsilon} = \frac{\varepsilon}{\frac{\rho_c \Delta x^5}{\Delta t^2}}$, $\tilde{\lambda}_0 = \frac{\lambda_0}{1}$, $\tilde{\boldsymbol{\Phi}} = \boldsymbol{\Phi}$, $\tilde{\mathbf{A}}^{-1} = \frac{\mathbf{A}^{-1}}{\Delta x^2}$, $\tilde{\mathbf{G}}^{-1} = \frac{\mathbf{G}^{-1}}{\frac{1}{\Delta x^2}}$, $\tilde{\mathbf{s}} = \frac{\mathbf{s}}{\Delta x}$

So

$$\tilde{\boldsymbol{\tau}}^{Ad} = -\frac{\varepsilon}{\frac{\rho_c \Delta x^5}{\Delta t^2}}\lambda_0\left(1-\lambda_0\right)^2 \Phi^{-\frac{1}{2}}\left[\frac{\mathbf{A}^{-1}}{\Delta x^2}\cdot\frac{\mathbf{G}^{-1}}{\frac{1}{\Delta x^2}}\cdot\frac{\mathbf{s}}{\Delta x}\right]\times\left[\frac{\mathbf{G}^{-1}}{\frac{1}{\Delta x^2}}\cdot\frac{\mathbf{s}}{\Delta x}\right] \tag{16}$$

$$= \frac{\mathbf{T}^{Ad}}{\frac{\rho_c \Delta x^5}{\Delta t^2}}$$

Thus there is no dimension problem with adhesion torque.

## Stokes resistance force constant:

As a force, the dimension of Stokes resistance force should be $\frac{\rho_c \Delta x^4}{\Delta t^2}$.

From its definition it is known that the dimension of viscosity $\mu$ is $\frac{\mu\cdot\Delta t}{\rho_c\cdot\Delta x^2}$ and the dimension of velocity $\mathbf{u}$ is $\frac{\Delta x}{\Delta t}$. Then I need to estimate the dimension of $\mathbf{K}^t$.

For $\mathbf{K}^t$ it is known following equations (which are discussed in Section 3.6, see Section 3.6 for physical interpretation of each parameter in equations)

$$\mathbf{K}^t = 16\pi\left(\frac{1}{\chi+a_1^2\alpha_1}\mathbf{u}_1\otimes\mathbf{u}_1+\frac{1}{\chi+a_2^2\alpha_2}\mathbf{u}_2\otimes\mathbf{u}_2+\frac{1}{\chi+a_3^2\alpha_3}\mathbf{u}_3\otimes\mathbf{u}_3\right) \tag{17}$$

$$\alpha_k = \int_0^\infty \frac{d\lambda}{\left(a_k^2+\lambda\right)\Delta\lambda} \quad (k=1,2,3) \tag{18}$$

$$\chi = \int_0^\infty \frac{d\lambda}{\Delta\lambda} \quad (k=1,2,3) \tag{19}$$

$$\Delta\lambda = \sqrt{\left(a_1^2+\lambda\right)\left(a_2^2+\lambda\right)\left(a_3^2+\lambda\right)} \tag{20}$$

From equation (215), dimension of $\lambda$ is $\Delta x^2$. Thus equation (214) turns to

$$\chi = \int_0^\infty \frac{d\lambda}{\sqrt{\left(a_1^2+\lambda\right)\left(a_2^2+\lambda\right)\left(a_3^2+\lambda\right)}} \tag{21}$$

Therefore

$$\begin{aligned}
\tilde{\chi} &= \int_0^\infty \frac{d\dfrac{\lambda}{\Delta x^2}}{\sqrt{\left(\dfrac{a_1^2}{\Delta x^2}+\dfrac{\lambda}{\Delta x^2}\right)\left(\dfrac{a_2^2}{\Delta x^2}+\dfrac{\lambda}{\Delta x^2}\right)\left(\dfrac{a_3^2}{\Delta x^2}+\dfrac{\lambda}{\Delta x^2}\right)}} \\
&= \int_0^\infty \frac{\dfrac{1}{\Delta x^2}d\lambda}{\dfrac{1}{\Delta x^3}\sqrt{\left(a_1^2+\lambda\right)\left(a_2^2+\lambda\right)\left(a_3^2+\lambda\right)}} \\
&= \frac{\displaystyle\int_0^\infty \frac{d\lambda}{\sqrt{\left(a_1^2+\lambda\right)\left(a_2^2+\lambda\right)\left(a_3^2+\lambda\right)}}}{\dfrac{1}{\Delta x}} \\
&= \frac{\chi}{\dfrac{1}{\Delta x}}
\end{aligned} \tag{22}$$

i.e. the dimension of $\chi$ is $\frac{1}{\Delta x}$.

It is known that $\tilde{a}_k = \frac{a_k}{\Delta x}\left(k=1,2,3\right)$ and $\tilde{\lambda} = \frac{\lambda}{\Delta x^2}$, so equation (213) turns to

$$\begin{aligned}
\tilde{\alpha}_k &= \int_0^\infty \frac{d\left(\dfrac{\lambda}{\Delta x^2}\right)}{\left(\dfrac{a_k^2}{\Delta x^2}+\dfrac{\lambda}{\Delta x^2}\right)\sqrt{\left(\dfrac{a_1^2}{\Delta x^2}+\dfrac{\lambda}{\Delta x^2}\right)\left(\dfrac{a_2^2}{\Delta x^2}+\dfrac{\lambda}{\Delta x^2}\right)\left(\dfrac{a_3^2}{\Delta x^2}+\dfrac{\lambda}{\Delta x^2}\right)}} \quad \left(k=1,2,3\right) \\
&= \int_0^\infty \frac{\dfrac{1}{\Delta x^2}d\left(\lambda\right)}{\dfrac{1}{\Delta x^2}\left(a_k^2+\lambda\right)\dfrac{1}{\Delta x^3}\sqrt{\left(a_1^2+\lambda\right)\left(a_2^2+\lambda\right)\left(a_3^2+\lambda\right)}} \\
&= \frac{\displaystyle\int_0^\infty \frac{d\left(\lambda\right)}{\left(a_k^2+\lambda\right)\sqrt{\left(a_1^2+\lambda\right)\left(a_2^2+\lambda\right)\left(a_3^2+\lambda\right)}}}{\dfrac{1}{\Delta x^3}} \\
&= \frac{\alpha_k}{\dfrac{1}{\Delta x^3}}
\end{aligned} \tag{23}$$

Thus the dimension of $\alpha_k \left(k=1,2,3\right)$ is $\frac{1}{\Delta x^3}$.

At last, with all the results above, equation (212) turns to

$$\begin{aligned}
\mathbf{K}^{\prime\prime} &= 16\pi\left(\frac{1}{\frac{\chi}{\frac{1}{\Delta x}}+\frac{a_1^2}{\Delta x^2}\cdot\frac{\alpha_1}{\frac{1}{\Delta x^3}}}\mathbf{u}_1\otimes\mathbf{u}_1+\frac{1}{\frac{\chi}{\frac{1}{\Delta x}}+\frac{a_2^2}{\Delta x^2}\cdot\frac{\alpha_2}{\frac{1}{\Delta x^3}}}\mathbf{u}_2\otimes\mathbf{u}_2+\frac{1}{\frac{\chi}{\frac{1}{\Delta x}}+\frac{a_3^2}{\Delta x^2}\cdot\frac{\alpha_3}{\frac{1}{\Delta x^3}}}\mathbf{u}_3\otimes\mathbf{u}_3\right)\\
&= 16\pi\left(\frac{1}{\Delta x\chi+\Delta x a_1^2\alpha_1}\mathbf{u}_1\otimes\mathbf{u}_1+\frac{1}{\Delta x\chi+\Delta x a_2^2\alpha_2}\mathbf{u}_2\otimes\mathbf{u}_2+\frac{1}{\Delta x\chi+\Delta x a_3^2\alpha_3}\mathbf{u}_3\otimes\mathbf{u}_3\right)\\
&= \frac{1}{\Delta x}\cdot 16\pi\left(\frac{1}{\chi+a_1^2\alpha_1}\mathbf{u}_1\otimes\mathbf{u}_1+\frac{1}{\chi+a_2^2\alpha_2}\mathbf{u}_2\otimes\mathbf{u}_2+\frac{1}{\chi+a_3^2\alpha_3}\mathbf{u}_3\otimes\mathbf{u}_3\right)\\
&= \frac{\mathbf{K}^t}{\Delta x}
\end{aligned}$$

(24)

Thus the dimension of $\mathbf{K}^t$ is $\Delta x$, then the dimension of $\left[\mathbf{K}^{\prime\prime}\right]^{-1}$ is $\frac{1}{\Delta x}$.

## Stokes resistance torque constant:

Similar to equation (212), equation (141) can be transformed as follows

$$\begin{aligned}
\mathbf{K}^{\prime\prime} &= 16\pi\left(\frac{\frac{a_2^2}{\Delta x^2}+\frac{a_3^2}{\Delta x^2}}{\frac{a_2^2}{\Delta x^2}\cdot\frac{\alpha_2}{\frac{1}{\Delta x^3}}+\frac{a_3^2}{\Delta x^2}\cdot\frac{\alpha_3}{\frac{1}{\Delta x^3}}}\mathbf{u}_1\otimes\mathbf{u}_1+\frac{\frac{a_1^2}{\Delta x^2}+\frac{a_3^2}{\Delta x^2}}{\frac{a_1^2}{\Delta x^2}\cdot\frac{\alpha_1}{\frac{1}{\Delta x^3}}+\frac{a_3^2}{\Delta x^2}\cdot\frac{\alpha_3}{\frac{1}{\Delta x^3}}}\mathbf{u}_2\otimes\mathbf{u}_2+\frac{\frac{a_2^2}{\Delta x^2}+\frac{a_1^2}{\Delta x^2}}{\frac{a_2^2}{\Delta x^2}\cdot\frac{\alpha_2}{\frac{1}{\Delta x^3}}+\frac{a_1^2}{\Delta x^2}\cdot\frac{\alpha_1}{\frac{1}{\Delta x^3}}}\mathbf{u}_3\otimes\mathbf{u}_3\right)\\
&= 16\pi\left(\frac{1}{\Delta x^3}\cdot\frac{a_2^2+a_3^2}{a_2^2\alpha_2+a_3^2\alpha_3}\mathbf{u}_1\otimes\mathbf{u}_1+\frac{1}{\Delta x^3}\cdot\frac{a_1^2+a_3^2}{a_1^2\alpha_1+a_3^2\alpha_3}\mathbf{u}_2\otimes\mathbf{u}_2+\frac{1}{\Delta x^3}\cdot\frac{a_2^2+a_1^2}{a_2^2\alpha_2+a_1^2\alpha_1}\mathbf{u}_3\otimes\mathbf{u}_3\right)\\
&= \frac{1}{\Delta x^3}\mathbf{K}^r
\end{aligned}$$

(25)

So the dimension of $\mathbf{K}^r$ is $\Delta x^3$.

## Velocity:

From its definition, it is known the dimension of velocity should be $\frac{\Delta x}{\Delta t}$. I have equation (147)

$$\mathbf{u} = \frac{1}{\mu}\left[\mathbf{K}^t\right]^{-1}\sum \mathbf{F} \tag{26}$$

I also know that $\left[\tilde{\mathbf{K}}^t\right]^{-1} = \frac{\left[\mathbf{K}^t\right]^{-1}}{\frac{1}{\Delta x}}$, $\tilde{\mathbf{F}}^{con} = \frac{\mathbf{F}^{con}}{\frac{\rho_c \Delta x^4}{\Delta t^2}}$, the dimension of viscosity $\mu$ is $\frac{\rho_c \Delta x^2}{\Delta t}$, thus

$$\begin{aligned}\tilde{\mathbf{u}}_0 &= \frac{1}{\frac{\mu}{\frac{\rho_c \Delta x^2}{\Delta t}}} \cdot \frac{\left[\mathbf{K}^t\right]^{-1}}{\frac{1}{\Delta x}} \cdot \frac{\sum \mathbf{F}\left(r_{ij}\right)}{\frac{\rho_c \Delta x^4}{\Delta t^2}} \\ &= \frac{1}{\frac{\Delta x}{\Delta t}} \cdot \frac{1}{\mu}\left[\mathbf{K}^t\right]^{-1}\sum \mathbf{F}\left(r_{ij}\right) \\ &= \frac{1}{\frac{\Delta x}{\Delta t}} \cdot \mathbf{u}_0\end{aligned} \tag{27}$$

Thus I can ensure that the velocity equation does not have dimension problem.

## Angular velocity:

The definition of angular velocity is the angle that the reference object turned in unit time [classical mechanics]. The angle is already dimensionless, or I can say its dimension is 1. Then the dimension of angle per unit time is $\frac{1}{\Delta t}$. From equation (149):

$$\boldsymbol{\omega} = \frac{1}{\mu}\left[\mathbf{K}^r\right]^{-1}\sum \mathbf{T} \tag{28}$$

I also know that $\tilde{\mu} = \frac{\mu}{\frac{\rho_c \Delta x^2}{\Delta t}}$, $\tilde{\mathbf{K}}^r = \frac{\mathbf{K}^r}{\Delta x^3}$, $\tilde{\mathbf{T}}^{Con} = \frac{\mathbf{T}^{Con}}{\frac{\rho_c \Delta x^5}{\Delta t^2}}$, so

$$\tilde{\boldsymbol{\omega}} = \frac{1}{\frac{\mu}{\frac{\rho_c \Delta x^2}{\Delta t}}} \cdot \frac{\left[\mathbf{K}^r\right]^{-1}}{\frac{1}{\Delta x^3}} \cdot \frac{\sum \mathbf{T}_{ij}}{\frac{\rho_c \Delta x^5}{\Delta t^2}} = \frac{\frac{1}{\mu}\left[\mathbf{K}^r\right]^{-1}\sum \mathbf{T}_{ij}}{\frac{1}{\Delta t}} = \frac{\boldsymbol{\omega}}{\frac{1}{\Delta t}} \tag{29}$$

Thus I can say that the angular velocity has no dimension problem in its equation.

Now I are sure there is no dimension problem with any equations or physical quantities I are using, then I can estimate the value of all the constants. The next step

is to estimate values of parameters in the model.

## Step 2: Value of parameters

Given satisfactory dimensional analysis for all the physical entities in my model, some parameters require constant values. Given a target system of cell-cell interactions in media, I can determine the physical interpretation of parameters and value ranges for these physical entities, and then from the range I may find proper value for each of them. Although my model is used for two purposes, there are common parts in the estimation of value of parameters, because for both purpose, the parameters should satisfy particular conditions, while the actual values of parameters may vary. Therefore in this section I firstly discuss the common conditions that parameters should satisfy for both purposes, and then calculate values of parameters separately. The values of parameters being used in simulation are listed in Section 4.1.3 and Section 6.3 (some values are tuned due to simulation result).

First, the unit length, unit time and unit density are the most basic entities, because all the other entities can be represented by combination of them. The physical parameters of the model (the unit time and length, and the constants involved in computing the forces) depend on the following values:
- The typical size of a cell;
- The range of ellipsoidal shapes a cell may adopt;
- The mean density of a cell;
- The dynamic viscosity of the fluid medium;
- The maximum speed at which a cell may move in the medium.

I begin with a basic force analysis, and when analysis deals with the above values, I work out the range of suitable values. Then the values for parameters are chosen according to requirements of each purpose. At last, the constants in contact force and adhesion force calculation are to be calculated based on other parameters. Note that although I do not consider the gravity in the model, I consider it for common condition that parameters should satisfy. In this process I derive the range of parameters so that the value of gravity does not significantly affect the condition.

Assume two non-identical ellipsoids and they do not contact each other. According to Newton's laws, $\mathbf{F} = m\mathbf{a}$, in which $m$ is mass of an ellipsoid, $\mathbf{F}$ is a summary of all the forces that affect on it, and $\mathbf{a}$ is the acceleration of an ellipsoid and equals $\frac{d\mathbf{v}}{dT}$, where $\mathbf{v}$ is the velocity of ellipsoid. I consider the gravity of ellipsoid for now, according to equation (147), the summary of all the forces is $m\mathbf{g} - \mu\mathbf{K}^t\mathbf{v}$, in which $m\mathbf{g}$ gravity of ellipsoidal agent is, and $-\mu\mathbf{K}^t\mathbf{v}$ is the Stokes resistance force. So

$$m\frac{d\mathbf{v}}{dT} = m\mathbf{g} - \mu\mathbf{K}^t\mathbf{v} \qquad (1)$$

Note that in this equation $dT$ is not unit time length $\Delta t$, $\Delta t$ is a fixed entity and has a value, while $dT$ means the change of velocity over time. Integral equation (201),

$$T = \frac{1}{-\frac{\mu \mathbf{K}^t}{m}} \cdot \ln(\mathbf{g} - \frac{\mu \mathbf{K}^t}{m}\mathbf{v}) + C \tag{2}$$

In which $C$ is constant. To determine the value of C I need to consider the initial condition. At the point that $T_0 = 0$, the ellipsoidal agents are planted into substrate so that the velocity $\mathbf{v}_0$ is a zero vector. Substituting this in to (202), I have $C = \mathbf{g}$. So that

$$\lim_{T \to +\infty} \mathbf{v} = \frac{m\mathbf{g}}{\mu \mathbf{K}^t} \tag{3}$$

Now I consider the approach that a cell being put into fluid, it falls into fluid but moves increasingly slowly until it eventually stops moving because of Stokes resistance. From equation (227) it is known that the terminal velocity approaches to zero when the effect of gravity is far smaller than $\mu \left| \mathbf{K}^t \right|$, i.e.

$$mg \ll \mu \left| \mathbf{K}^t \right| \tag{4}$$

To simplify the situation, I considering cells as a sphere, this will not affect my estimation. For example, for ellipsoid $\mathbf{A}$ with semi-radii $a_1$, $a_2$ and $a_3$ which satisfy $a_1 \le a_2 \le a_3$, I simplify the ellipsoid as a sphere with radius of the longest semi-radius $a_3$, then the volume of sphere should be larger than the ellipsoid thus the mass of sphere is larger than ellipsoid as well (assume they have the same density). If the mass of sphere satisfy condition $m_{sphere} g \ll \mu \left| \mathbf{K}^t \right|$, the mass of ellipsoid satisfy $m_{ellipsoid} g < m_{sphere} g \ll \mu \left| \mathbf{K}^t \right|$, which also satisfy the condition.

Then the mass of cell can be presented as $m = \rho_e \cdot \frac{4\pi r^3}{3}$, and Stokes resistance force constant $\left| \mathbf{K}^t \right| = 6\pi r$, so that $\frac{4\pi r^3 \rho_e}{3} \cdot g \ll 6\pi r \cdot \mu$. I let the dynamic viscosity $\mu$ have the value of $1 \times 10^{-3} Pa \cdot s$ (this is dynamic viscosity of water), gravity acceleration $g$ have the value of $10 m/s^2$, plus density of cell cannot be smaller than this value of water $\rho_c = 1000 kg/m^3$. With all these values I have an upper range of cell radius $r$: $r \ll 6.7 \times 10^{-4} m$.

In following section I use the density of water as unit density. Now the terminal speed

$$v_{term} = \frac{2}{9} \cdot \frac{\rho_e g}{\mu} \cdot r^2 \tag{5}$$

In which $\mu$ is dynamic viscosity, $r$ is radius of ellipsoid (simplified to a sphere). From this equation it is known that if the terminal velocity is small, the density of cells should be small, and radius of cells should also be small. At the same time the dynamic viscosity $\mu$ should be relatively large so that the terminal speed can be small.

Assuming the non-dimensional terminal speed is $\tilde{v}_{term} = \frac{v_{term}}{\Delta v}$, then according to the definition of speed, the unit speed equals unit length divided by unit time, i.e. $\Delta v = \frac{\Delta x}{\Delta T}$. Note that $\Delta T$ is not $\Delta t$, because $\Delta t$ is a concept used for dimension analysis only, while $\Delta T$ is the unit time in my model.

Then I will represent the radius in equation (228) by unit entities. It is straight forward to set $\Delta x = r$, i.e. the length of unit length is equal to the typical radius of cell. It is also possible to let typical radius equals several times of unit length to keep potential of representing an entity in between.

For parameters in Chapter 4, in vascular formation, endothelial cells show strong anisotropy. Thus I want a long and thin shape for the agents, so that when other agents approach from different directions, the interaction is significantly different. I assume $\tilde{r} = \frac{r}{\Delta x} \in [1, \quad 4]$ at first, later I can choose a proper value from this range.

Obviously the density of cell is bigger than water, but not too much bigger otherwise the terminal speed will increase. Assume the density of cell $\rho_e = 2\rho_c = 2000 kg/m^3$

Substitute all these to equation (228), and then it turns to

$$\Delta x \Delta T \in \left[1.406\times10^{-8}\cdot\tilde{v}_{term}, \quad 2.25\times10^{-7}\cdot\tilde{v}_{term}\right] \tag{6}$$

I take typical radius of cell as $r = 5\times10^{-6} m$, so that $\Delta x \in \left[1.25\times10^{-6}, \quad 5\times10^{-6}\right] m$, and $\Delta T \in \left[2.812\times10^{-3}\cdot\tilde{v}_{term}, \quad 1.8\times10^{-1}\cdot\tilde{v}_{term}\right]$. Let $\Delta x$ be $1.25\times10^{-6}$ m, then $v_{term} = 1.11\times10^{-4} m/s$, and $\Delta T \in \left[2.4971\times10^{-1}, \quad 1.5984\times10^{1}\right] s$. Let $\Delta T$ be 1s, so $\Delta v = \frac{\Delta x}{\Delta T} = 1.25\times10^{-6} m/s$.

Parameter values derived so far:
Unit length $\Delta x = 1.25\times10^{-6} m$;
Unit time step $\Delta T = 1s$;
Unit density $\rho_c = 1000 kg/m^3$;
Unit speed $\Delta v = 1.25\times10^{-6} m/s$.

I next need to estimate the value of constant $K$ in contact force/torque equations; and $\varepsilon$ in adhesion energy, force and torque equations. Note that I need non-dimensional entities so I need to estimate their value then divide by their dimensions.

As the Young's' modulus for cell is $0.471\pm0.081$ kPa (Xiaofang Cai *et al*. 2009), so for two same cells, $E_1 = E_2 = (0.471\pm0.081)\times10^3$ Pa. And $\nu_1 = \nu_2 = \frac{1}{3}$, so

$$K = \frac{3}{4}E = 0.35325\times10^3 \tag{7}$$

Because $\tilde{K} = \frac{K\Delta T^2}{\rho_c \cdot \Delta x^2}$,

$$\tilde{K} = 2.2608 \times 10^{11} \tag{8}$$

Assume ellipsoid $A$ and $B$ are spheres with radius of $r$.

$$\varepsilon = K\left[r + r\right]^2 r \cdot \left(1 - \sqrt{\Phi}\right)^{\frac{3}{2}} \tag{9}$$

$$\frac{\varepsilon}{K} = 4r^3 \cdot \left(1 - \sqrt{\Phi}\right)^{\frac{3}{2}} \tag{10}$$

From (234), it is known that the bigger $\Phi$ is, the smaller $\frac{\varepsilon}{K}$ is. And

$$\underset{\Phi \to 1}{Lim}\left(\frac{\varepsilon}{K}\right) = 0 \tag{11}$$

Now I assume the adhesion force balances with contact force when contact potential is 0.9216 (which is square of 0.96). Then

$$\varepsilon = K\left[r + r\right]^2 r \cdot \left(1 - 0.96\right)^{\frac{3}{2}} \tag{12}$$

$$\frac{\varepsilon}{K} = 4 \cdot \left(5 \times 10^{-6}\right)^3 \cdot \left(0.2\right)^3 = 4 \times 10^{-18} \tag{13}$$

So that

$$\varepsilon = 1.413 \times 10^{-15} \tag{14}$$

As I use dimensional value of parameters in model

$$\text{%} = \frac{\varepsilon \Delta T^2}{\rho_c \Delta x^5} = 4.6301 \times 10^{11} \tag{15}$$

For parameters in Chapter 6, due to the interval of two time-lapse images are taken is 10 minuets, and in my simulation I output the position of agents every 10 simulation loops (See Figure 4.1.3.1 for detail), I let $\Delta T$ equal to 1 minute. I assume the density of cancer cells is not significantly different from endothelial cells, and the viscosity of environment is also nearly the same. Then the derivation about terminal speed remains same,

$$v_{term} = 1.11 \times 10^{-4} m/s \tag{16}$$

I assume the value of unit speed remains the same, i.e.

$$\Delta v = \frac{\Delta x}{\Delta T} = 1.25 \times 10^{-6} m/s \tag{17}$$

And then $\Delta x = \Delta v \cdot \Delta T = 7.5 \times 10^{-5}$ m. From time-lapse images of control group, I see that the cancer cells round circles from a top-down view; the radius of circle is about $5.0 \times 10^{-6}$ m. Thus I set three semi-radii of agent as $1.25 \times 10^{-6}$ m, $5.0 \times 10^{-6}$ m, $5.0 \times 10^{-6}$ m, so that the agent is a thin slice sticking on substrate.

I also assume the Young's modulus of cancer cell is not significantly different from endothelial cell. Thus for cancer cell I also have

$$K = \frac{3}{4}E = 0.35325 \times 10^{3} \tag{18}$$

And because $\tilde{K} = \dfrac{K\Delta T^2}{\rho_c \cdot \Delta x^2}$

$$\tilde{K} = 1.1304 \times 10^{11} \tag{19}$$

Because of equation (237), $\dfrac{\varepsilon}{K} = 4 \times 10^{-18}$, thus

$$\varepsilon = 1.413 \times 10^{-15} \tag{20}$$

Then from $\tilde{\varepsilon} = \dfrac{\varepsilon \Delta T^2}{\rho_c \Delta x^5}$

$$\tilde{\varepsilon} = 1.072 \times 10^{15} \tag{21}$$

# Appendix B: Code Structure and Simulation Issue

## Simulation class

| | |
|---|---|
| class Ellipsoid; | Contains all the basic physical aspect of model, including all the force/torque/velocity/angular velocity that mentioned above. |
| class Cell; | Derived from Ellipsoid class. And the biology functions are placed here. For now there is no biology feature but it will be added in future. |
| class Simulation; | Contains an array of object of class Cell as data member and control the simulation such as how long it will run or in which file the data result will be. |

The contact force and contact torque describe the physical interaction between cells, so that there is an Ellipsoid class to contain all physical-relative functionality. Also the Ellipsoid class contains the method to calculate velocity and angular velocity.

Because my aim is to simulate the behaviour of cell, there are forces and other factors that affect it. The Cell class is derived from Ellipsoid class, so that it contains the physical functions as well. The Cell class also contains cell cycle and some cell related forces and torques. Now they are kept blank, because it makes it clearer to test the physical part of the model first. Now the code is a prototype of the final model, each feature will be added in after the previous step is proved. When I operate an object of Cell class, which is what I do in the simulation, the physical work is done by functions of its parent class, the biological work is done by Cell functions (they are blank at the moment though).

At last, a list of the Cell object is included in the Simulation class. The Simulation class reads settings from file, initializes the cell list, and controls all the calculation. It also saves the position and direction of each cell regularly.

# Cell class

Class HippoCell

**Data**

-Firstly HippoCell class has a mark to tell which stage it is in:

CellStage currentCellStage;

The CellStage is an enum type which is defined in the header file.

-HippoCell class has a group of timers which controls how long each stage should be:

int GrowthStageLength; // how long it will take for the cell to grow to full size and turn to split //stage

int SplitStageLength; // how long the split stage should be

int MaxAge; // if cell cycle >= this number and cell not in split stage, the cell gets to die out stage.

Each of these 3 timers is generated from its range which is combined by fixed max value and min value, which is defined in header file. Here I need an algorithm to have normal distribution of stage length.

-HippoCell class also has 3 tuners for each timer; to represent the effect that environment does on cells.

int GrowthStageTuner;

int SplitStageTuner;

int MaxAgeTuner;

These can be dynamic values that changes along the environment.

-There should be a percentage value. When cell grows to full size, it has a chance to split or die. This value is a decimal between 0 and 1, and should be different for different cell types.

float SplitChance;

-In the simulation there will be two types of cells, which means I will have two lists of cells, each one belongs to a cell type and has its own physical parameter set.

-There also should be a mark to tell if the cell is affected by Hippo signal or not.

-If the cell is affected by Hippo signal, it should has 2 counters, one to represent how many contacted neighbours it has and the other to represent the average/total contact potential.

int ContactedNeighbor;

float AveragePotential; // only with contacted neighbors

**Methods**

- As a class derived from class Cell, HippoCell class contains all the methods that Cell class has. However it also has special methods related to stage control and fate determination.

void SetCellStage( CellStage stage );

CellStage GetCellStage();

void FateDetermination(); // decide if cell should turns to any stage

There also should be a method for splitting, in which a new HippoCell object is built and added to cell list.

-However I need a group of methods so that I can manually set the length of each stage:

void SetGrowthStage();

void SplitStage();

void SetMaxAge();

# Single growth control

**Code part reflecting growth curve**

The parameters relative with growth curve are reproduce rate of single cell, saturation density of environment (these 2 controls the population growth curve); and information about single cell volume during the cell cycle, which controls the single cell growth curve.

To store the single cell growth curve information, I have a new structure called

SingleCellVolume, which contains the maximum age of cell and the start and end time of 2 slops. It should look like this:

```
Struct SingleCellVolume
{
    int m_MaxAge;
    int m_GrowthtStartTime;
    int m_GrowthEndTime;
    float m_GrowthCurveSlope;
 };
```

The SingleCellVolume is stored in class Simulation as an array, and initialize from .txt file. When it is time to build each cell, its value is assigned into cell as well. I keep a copy in class Simulation so that I can check how each kind of cell grows at any time during the simulation.

The saturation density is stored in class Simulation, because it is part of simulation environment. The reproduce rate, however, may differ for each type of cells, so that it is stored in each cell.

The parameters of growth curve of single and group cells are loaded in a new function called initGrowthPara(…) called from function SimulationPreStart(…) of class Simulation. The function initGrowthPara(…) reads the growth curve of single cell and reproduce rate from .txt file. The cell related data is stored

**Cell age distribution**

In the experiment, cells will not split at exact time, their age abbeys the normal distribution. To implement this, I need to define the desired value (typical value) and variance of cell age. Similarly, each phrase of cell requires a typical value and variance to define.

To make cells obey the normal distribution, in program, I generate random numbers that obey normal distribution.

# TBB template

This topic is discussed in paper for research student 2012. Parallel_do is the simplest template to do the same operation to a list of data of foxed size. The data list is divided into smaller segments and distributed to processing threads to process concurrently. The number of concurrent thread can be set by the user or left for the library to decide. Instead of processing data serially parallel_do can process several data at the same time, and in this way the efficiency is expected to be significantly improved. Further, as the division may happen on any part of the data list, this template requires the data list to be random accessible so that each threads is able to find the data segment it is assigned to operate. parallel_reduce is another template used in the simulation. Its task is to apply the reduction operation to reconstitute the data list following parallel execution, i.e. to aggregate the separate calculations, waiting for all to complete, to re-integrate the data list. For the same reason parallel_reduce requires random access as well.

In the simulation, these two templates are combined together to achieve maximum improvement of efficiency: the parallel_reduce is used to sum up the force between one cell and all other cells, and parallel_do is in charge of do this operation to each cell in the cell list. During the period in which the model parameters are tuned to be realistic, the simulation is run on a machine with Intel quad-core CPU Q8300 2.50G Hz.

(From research student 2012) I measure how much the TBB parallel technique speeds up the simulation with 2800 cells, which is a typical number to show vessel pattern. Firstly the running time of sequence code is recorded, as the following chart shows, the sequence code uses 9216 ms in each loop of all cells. An optimally parallel code will use 1.00, 0.53, 0.38, and 0.30 times of sequence code time (see Appendix for detail), by running on one, two, three and four CPUs, respectively. Thus the ideal looping times are 9216.00 ms, 4876.77 ms, 3460.36 ms and 2752.15 ms. Once the code is paralleled by one, two, three and four CPUs the observed running time for each loop is 13328.00 ms, 8625.00 ms, 6079.00 ms, 5007.50 ms, which is 1.46, 0.95, 0.67 and 0.55 times of sequence time. And compared to parallel with one thread,

| Number of Cores | Ideal Speedup | Sequence code Time (ms) | Ideal Running Time (ms) | Parallel Code Time (ms) | Parallel vs. Sequence | Actual Speedup |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 9126.00 | 9126.00 | 13328.00 | 1.46 | 1.00 |
| 2 | 0.53 | 9126.00 | 4876.77 | 8625.00 | 0.95 | 0.65 |
| 3 | 0.38 | 9126.00 | 3460.36 | 6079.00 | 0.67 | 0.46 |
| 4 | 0.30 | 9126.00 | 2752.15 | 5007.50 | 0.55 | 0.38 |

Sequence and Parallel code running time and speedup

(From research student 2012) It can be seem more clearly with a chart. As the followed one shows, the purple column represents the time spent in sequential execution, the dark red column shows the ideal time spent in parallelised code, and the yellow column represents the actual time spent in parallelised code. I can see that parallel with one thread uses significantly more time than sequential code. This is because building up threads costs extra time, further more, the parallel code cannot have benefit from compiler optimization. However parallel with two threads reduce the looping time to almost same level of sequential code. And more CPUs make improvement to the running efficiency, three CPUs reduce running time to about 2/3, and four CPUs used about half of sequence time.

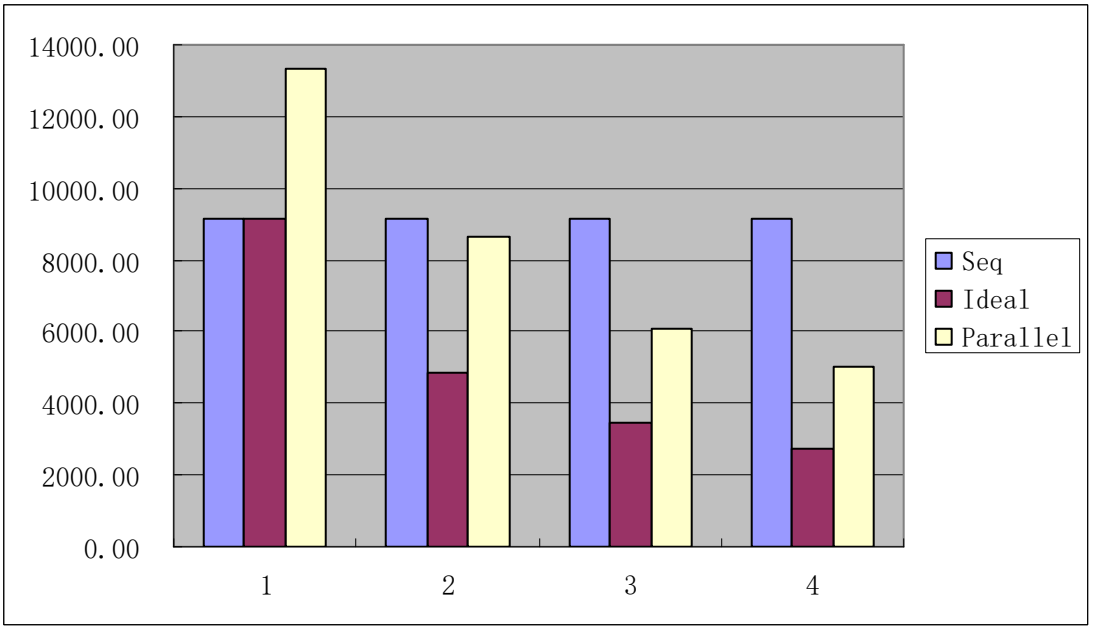



Chart 3.2 Parallel Efficiency with various numbers of CPUs

(From research student 2012) From the profiler it is known that the code spent 335 ms to write data to file, 1904 ms in the main loop and 30322 ms per loop. All these data are obtained by VS2010 Ultimate Profiler with sequence version of code. So that ideally, for N number of CPUs, the paralleled time should be $t(N) = 335 + 1904 + \frac{30322}{N}$ ms. And substitute $N = 1, 2, 3, 4$ I can get ideal running time, which are 32561 ms, 17400 ms, 12346.33 ms and 9819.35 ms, respectively. And then divide them by 32561 ms which is the total time spent in that profile analysis I get how many times the ideal parallel code should be faster.

# Appendix C: Values of Module 'IdentifyPrimaryObjects' in Cell Recognisation

In Section 6.1.2, the image of experiments are imported to CellProfiler to be identified. The main process of identification is done in the module 'IdentifyPrimaryObjects'. The interface of module 'IdentifyPrimaryObjects' is shown in Figure 6.1.2.1.

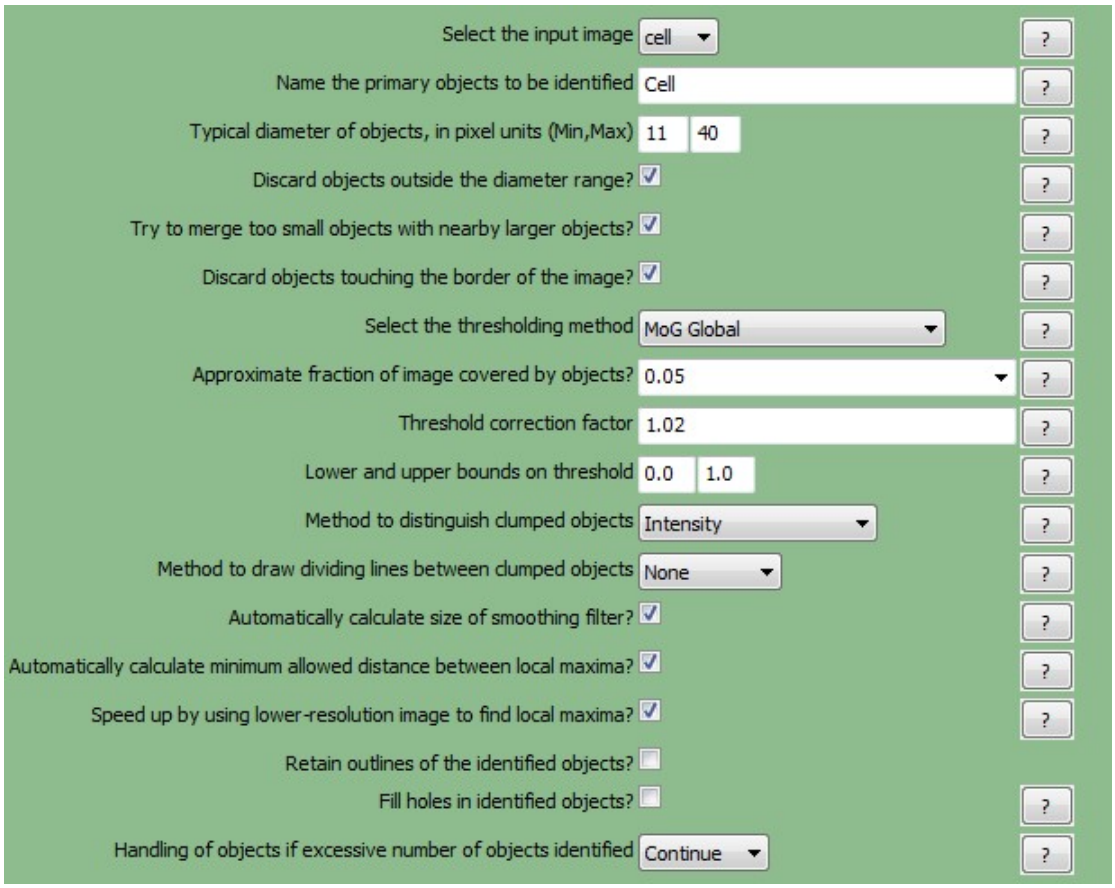



Figure 6.1.2.1 Interfact of module 'IdentifyPrimaryObjects'

As shown in Figure 6.1.2.1, the name of group of images set in 'NamesandTypes' is

passed to the first setting. Here the name of the group is ‘cell’. An important parameter is the range of diameter of objects, which is measured in number of pixels. From time-lapse images, it can be observed that the diameter of cells is around 10 pixels (when the shape of cell is round), so I set the lower range of diameter of object to 9, 10 or 11. The actual value may vary among images from the same experiment. The actual value is chosen to have most cells recognised while lease dust and other sundries recognised as cell. The higher range of diameter is set to 40, because on some of the images I can see the shape of cell is slender and 40 is large enough to let those cells to be recognized. These values provide the upper and lower limits of cell sizes that are necessary for cell recognition.

Another parameter that can significantly affect the result of recognition is the ‘Threshold correction factor’. If a value < 1.0 is entered, the threshold is adjusted to be more lenient so that more cells are recognised; on the contrary, a value > 1.0 means the threshold is adjusted more strictly (CellProfiler document). This value highly depends on a particular image and can also vary among images from the same experiment. The actual values used are shown in the following paragraphs.

The last key parameter is ‘Method to distinguish clumped objects’. CellProfiler provides four choices to be selected from. For the time-lapse image, the ‘’Intensity’ or ‘Shape’ is selected according to achieve better recognition.

The setting of ‘IdentifyPrimaryObjects’ is shown as follows in Figure 6.1.2.3.

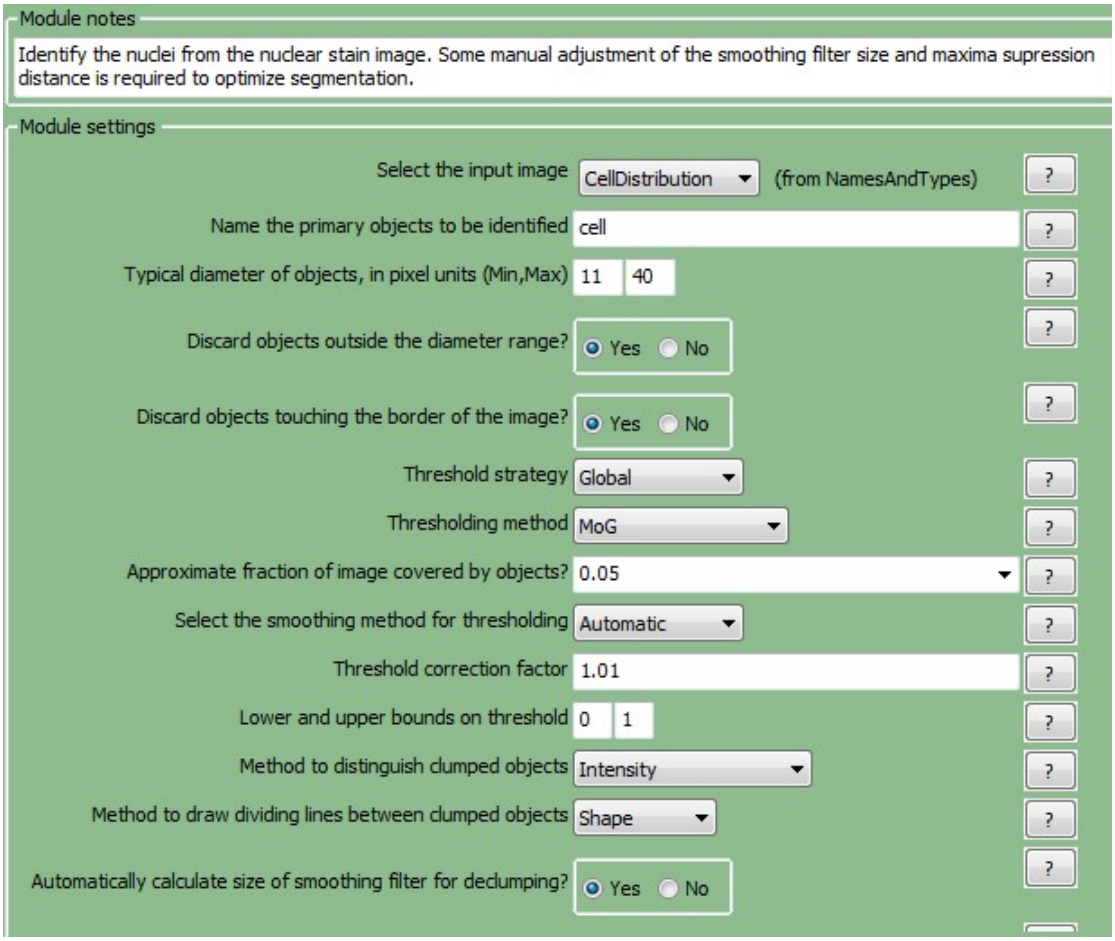


Figure 6.1.2.3 setting of 'IdentifyPrimaryObjects' to identify cells from time-lapse images

However for some images the threshold correction factor is modified to produce a better result. For example in HCT wild type hypoxia group, for image normhpf_0000.png, normhpf_0009.png and normhpf_0019.png, the following setting is used, in which the threshold correction factor is increased from 1.0 to 1.02.

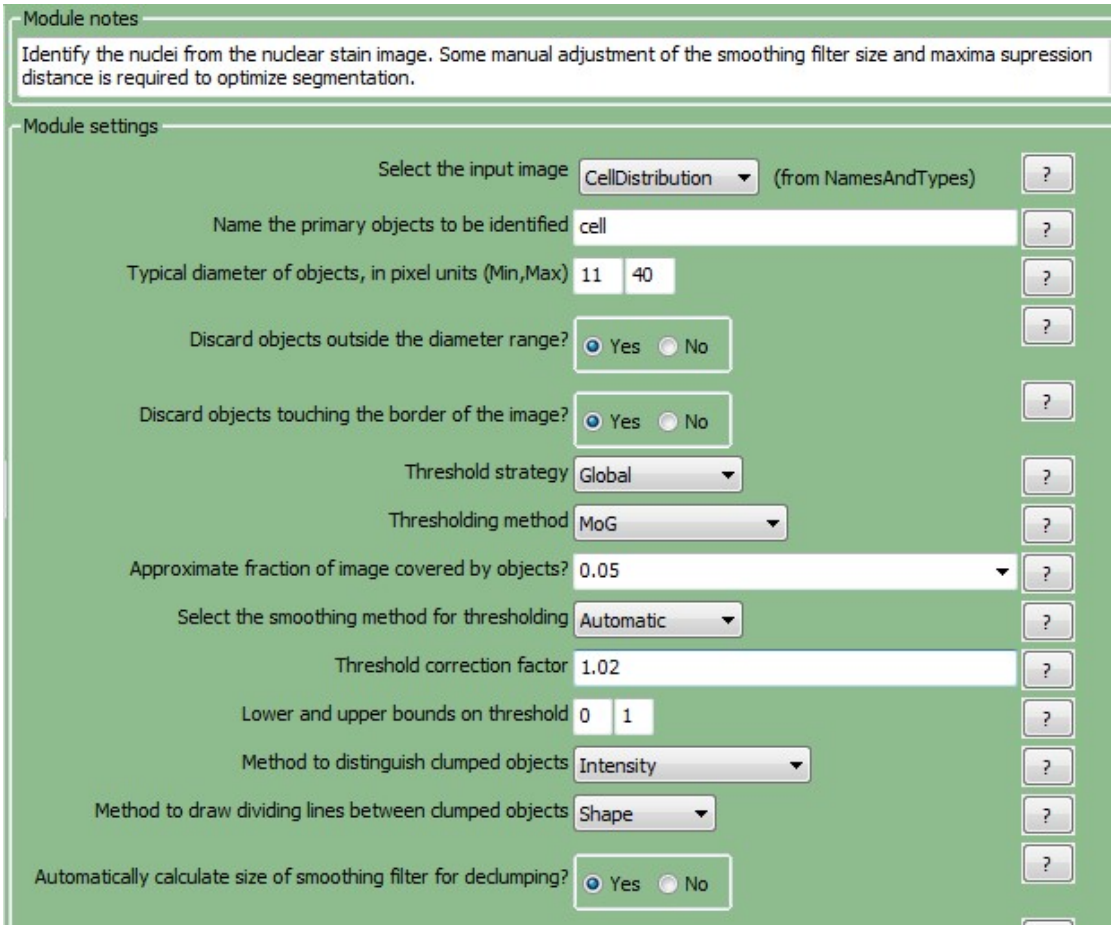



Figure 6.1.2.4 another setting of 'IdentifyPrimaryObjects'

Values of Threshold correction factor for all 13 images are shown in Table 6.1.2.1. The variation of threshold correction factor suggests that there are still slight differences among this group of images.

| Name of image | Value of Threshold correction factor |
|---|---|
| normhpf_0000.png | 1.02 |
| normhpf_0009.png | 1.02 |
| normhpf_0019.png | 1.02 |
| normhpf_0029.png | 1.05 |
| normhpf_0039.png | 1.05 |
| normhpf_0049.png | 1.03 |
| normhpf_0059.png | 1.03 |
| normhpf_0069.png | 1.03 |

| | |
|---|---|
| normhpf_0079.png | 1.03 |
| normhpf_0089.png | 1.03 |
| normhpf_0099.png | 1.03 |
| normhpf_0109.png | 1.03 |
| normhpf_0119.png | 1.03 |

Table 6.1.2.1 Values of Threshold correction factor for 13 images of the hypoxic group

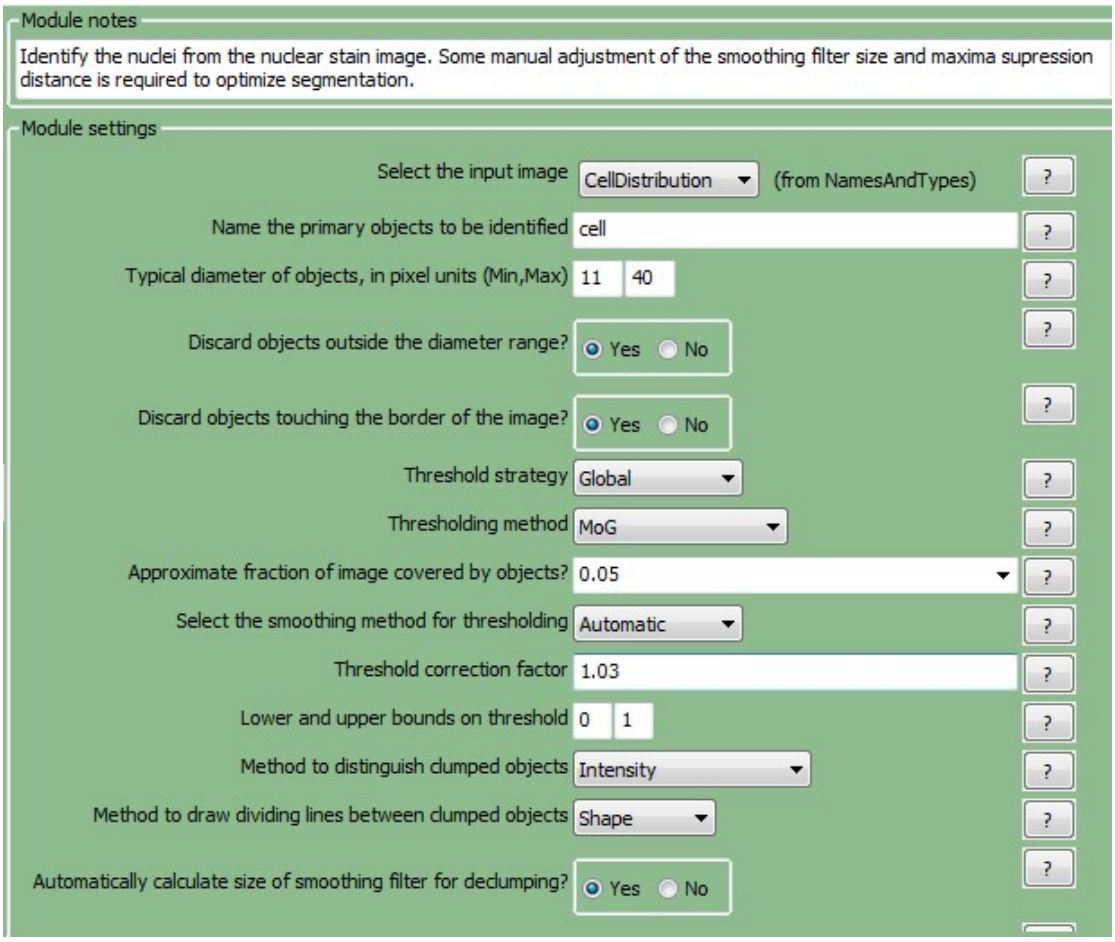



Figure 6.1.2.5 another setting of 'IdentifyPrimaryObjects'

Although the change of threshold correction factor seems to be small in value, the edge reorganization is sensitive to this factor. As Figure 6.1.2.5 shows, by increasing or decreasing this factor by 0.01, the outcome is completely different.

n 5FU+Hypoxia group, the variation of threshold correction factor is also found, as Table 6.1.2.2 shows.

| Name of image | Value of Threshold correction factor |
|---|---|
| normhpf_0000.png | 1.02 |
| normhpf_0010.png | 1.08 |
| normhpf_0020.png | 1.05 |
| normhpf_0030.png | 1.05 |
| normhpf_0040.png | 1.05 |
| normhpf_0050.png | 1.05 |
| normhpf_0060.png | 1.05 |
| normhpf_0070.png | 1.04 |
| normhpf_0080.png | 1.04 |

Table 6.1.2.2 Values of Threshold correction factor for 9 images of 5FU+Hypoxia group

In 5FU group, the variation of threshold correction factor is shown in Table 6.1.2.3.

| Name of image | Value of Threshold correction factor |
|---|---|
| normhpf_0000.png | 0.85 |
| normhpf_0010.png | 0.85 |
| normhpf_0020.png | 0.85 |
| normhpf_0030.png | 0.85 |
| normhpf_0040.png | 0.85 |
| normhpf_0050.png | 0.85 |

| | |
|---|---|
| normhpf_0060.png | 0.85 |
| normhpf_0070.png | 0.85 |
| normhpf_0080.png | 0.85 |
| normhpf_0090.png | 1.43 |
| normhpf_01000.png | 1.43 |
| normhpf_0110.png | 1.43 |

Table 6.1.2.3 Values of Threshold correction factor for 12 images of 5FU group

# Reference


Agnew, D. J. G., Green, J. E. F., Brown, T. M., Simpson, M. J., & Binder, B. J. (2014). Distinguishing between mechanisms of cell aggregation using pair-correlation functions. Journal of theoretical biology, 352, 16-23.

Allen, M. P. & Germano, G. (2002). Rigid body potentials, forces and torques

Allen, M. P. (2004). Introduction to molecular dynamics simulation. Computational Soft Matter: From Synthetic Polymers to Proteins, 23, 1-28.

Ambrosi, D., Gamba, A., & Serini, G. (2004). Cell directional and chemotaxis in vascular morphogenesis. Bulletin of mathematical biology, 66(6), 1851-1873.

Anderson, A., & Rejniak, K. (Eds.). (2007). Single-cell-based models in biology and medicine. Springer Science & Business Media.

Anderson, P. W. (1972). More is different. Science, 177(4047), 393-396.

Aschacher, T., Wolf, B., Enzmann, F., Kienzl, P., Messner, B., Sampl, S., ... & Bergmann, M. (2015). LINE-1 induces hTERT and ensures telomere maintenance in tumour cell lines. Oncogene.

Baddeley, A. (2008). Analysing spatial point patterns in R. Technical report, CSIRO, 2010. Version 4. Available at www. csiro. au/resources/pf16h. html.

Barenblatt, G. I. (1996). *Scaling, self-similarity, and intermediate asymptotics: dimensional analysis and intermediate asymptotics* (Vol. 14). Cambridge University Press.

Barleon, B., Sozzani, S., Zhou, D., Weich, H. A., Mantovani, A., & Marme, D. (1996). Migration of human monocytes in response to vascular endothelial growth factor (VEGF) is mediated via the VEGF receptor flt-1. *Blood*, *87*(8), 3336-3343.

Bergers, G., & Hanahan, D. (2008). Modes of resistance to anti-angiogenic therapy. *Nature Reviews Cancer*, *8*(8), 592-603.

Berne, B. J., & Pechukas, P. (1972). Gaussian model potentials for molecular interactions. *The Journal of Chemical Physics*, *56*(8), 4213-4216.

Beysens, D. A., Forgacs, G., & Glazier, J. A. (2000). Cell sorting is analogous to phase ordering in fluids. *Proceedings of the National Academy of Sciences*, *97*(17), 9467-9471.

Bhaskar, R., & Nigam, A. (1990). Qualitative physics using dimensional analysis. *Artificial Intelligence*, *45*(1), 73-111.

Bicknell, G. R., Snowden, R. T., & Cohen, G. M. (1994). Formation of high molecular mass DNA fragments is a marker of apoptosis in the human leukaemic cell line, U937. *Journal of cell science*, *107*(9), 2483-2489.

Binder, B. J., & Simpson, M. J. (2013). Quantifying spatial structure in experimental observations and agent-based simulations using pair-correlation functions. Physical Review E, 88(2), 022705.

Bivand, R. S., Pebesma, E. J., Gomez-Rubio, V., & Pebesma, E. J. (2008). Applied spatial data analysis with R (Vol. 747248717). New York: Springer.

Bodnar, A. G., Ouellette, M., Frolkis, M., Holt, S. E., Chiu, C. P., Morin, G. B., ... & Wright, W. E. (1998). Extension of life-span by introduction of telomerase into normal human cells. Science, 279(5349), 349-352.

Bown, J., Deeni, Y., Savage, A., Sampson, A., Khalil, H. S., Idowu, M., Zhelev, N. &

Li, Y. EBTNA Book (draft).
Bown, J., S Andrews, P., Deeni, Y., Goltsov, A., Idowu, M., AC Polack, Fiona AC Polack, Adam T Sampson, Mark Shovman, and Stepney, S. (2012). Engineering simulations for cancer systems biology. *Current drug targets*, *13*(12), 1560-1574.
Brenner, H. (1963). The Stokes resistance of an arbitrary particle. *Chemical Engineering Science*, *18*(1), 1-25.
Brenner, H. (1967). Coupling between the translational and rotational brownian motions of rigid particles of arbitrary shape: II. General theory. *Journal of colloid and interface science*, *23*(3), 407-436.
Brenner, H., & Gajdos, L. J. (1981). London-van der Waals forces and torques exerted on an ellipsoidal particle by a nearby semi-infinite slab. *Canadian Journal of Chemistry*, *59*(13), 2004-2018.
Brú, A., Albertos, S., Subiza, J. L., García-Asenjo, J. L., & Brú, I. (2003). The universal dynamics of tumor growth. *Biophysical journal*, *85*(5), 2948-2961.
Cai, X., Cai, J., Dong, S., Deng, H., & Hu, M. (2009). [Morphology and mechanical properties of normal lymphocyte and Jurkat revealed by atomic force microscopy]. *Sheng wu gong cheng xue bao= Chinese journal of biotechnology*, *25*(7), 1107-1112.
Canel, M., Serrels, A., Frame, M. C., & Brunton, V. G. (2013). E-cadherin–integrin crosstalk in cancer invasion and metastasis. Journal of cell science, 126(2), 393-401.
Cavalli, F., Gamba, A., Naldi, G., Semplice, M., Valdembri, D., & Serini, G. (2007). 3D simulations of early blood vessel formation. *Journal of Computational Physics*, *225*(2), 2283-2300.
Chauhan, V. P., Stylianopoulos, T., Martin, J. D., Popović, Z., Chen, O., Kamoun, W. S., ... & Jain, R. K. (2012). Normalization of tumour blood vessels improves the delivery of nanomedicines in a size-dependent manner. Nature nanotechnology, 7(6), 383-388.
Chen, Y., Cairns, R., Papandreou, I., Koong, A., & Denko, N. C. (2009). Oxygen consumption can regulate the growth of tumors, a new perspective on the Warburg effect. *PLoS One*, *4*(9), e7033.
Chenevert, T. L., McKeever, P. E., & Ross, B. D. (1997). Monitoring early response of experimental brain tumors to therapy using diffusion magnetic resonance imaging. Clinical Cancer Research, 3(9), 1457-1466.
Choudhary, B., Hanski, M. L., Zeitz, M., & Hanski, C. (2012). Proliferation rate but not mismatch repair affects the long-term response of colon carcinoma cells to 5FU treatment. *Cancer letters*, *320*(1), 56-64.
Cooper, G. M. & Hausman, R. E. (2013), The Cell: A Molecular Approach. 6th Edition. 2013. Sinauer Associates: Sunderland, MA.
Culver, C., Melvin, A., Mudie, S., & Rocha, S. (2011). HIF-1α depletion results in SP1-mediated cell cycle disruption and alters the cellular response to chemotherapeutic drugs. *Cell Cycle*, *10*(8), 1249-1260.
Daley, D. J., & Vere-Jones, D. (2003). An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods. Springer Science & Business Media.
Dallon, J. C. (2007). Models with lattice-free center-based cells interacting with continuum environment variables. In *Single-Cell-Based Models in Biology and Medicine* (pp. 197-219). Birkhäuser Basel.
David O. Morgan. (2007). The cell cycle, principles of control. New Science Press, 2007.
Deason-Towne, F., Perraud, A. L., & Schmitz, C. (2011). The Mg 2+ transporter MagT1 partially rescues cell growth and Mg 2+ uptake in cells lacking the

channel-kinase TRPM7. *FEBS letters*, *585*(14), 2275-2278.
Derjaguin, B. V., Muller, V. M., & Toporov, Y. P. (1975). Effect of contact deformations on the adhesion of particles. *Journal of Colloid and interface science*, *53*(2), 314-326.
Deutsch, A. (2007). Lattice-gas cellular automaton modeling of developing cell systems. In *Single-Cell-Based Models in Biology and Medicine* (pp. 29-51). Birkhäuser Basel.
Drasdo, D. (2007). Center-based single-cell models: An approach to multi-cellular organization based on a conceptual analogy to colloidal particles. In *Single-Cell-Based Models in Biology and Medicine* (pp. 171-196). Birkhäuser Basel.
Drasdo, D., & Hoehme, S. (2005). A single-cell-based model of tumor growth in vitro: monolayers and spheroids. *Physical biology*, *2*(3), 133.
Eckel, S., Brachat-Schwarz, W., Fleischer, F., Grabarnik, P., Schmidt, V. & Walla, W. Analysis of spatial correlations in marked point processes (power point)
Elmore, S. (2007). Apoptosis: a review of programmed cell death. Toxicologic pathology, 35(4), 495-516.
Faratian, D., Clyde, R. G., Crawford, J. W., & Harrison, D. J. (2009). Systems pathology—taking molecular pathology into a new dimension. *Nature Reviews Clinical Oncology*, *6*(8), 455-464.
Friedl, P., & Alexander, S. (2011). Cancer invasion and the microenvironment: plasticity and reciprocity. Cell, 147(5), 992-1009.
Gamba, A., Ambrosi, D., Coniglio, A., De Candia, A., Di Talia, S., Giraudo, E., ... & Bussolino, F. (2003). Percolation, morphogenesis, and Burgers dynamics in blood vessels formation. *Physical review letters*, *90*(11), 118101.
Giesen, C., Wang, H. A., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., ... & Bodenmiller, B. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature methods*, *11*(4), 417-422.
Goel, S., Fukumura, D., & Jain, R. K. (2012). Normalization of the tumor vasculature through oncogenic inhibition: An emerging paradigm in tumor biology. Proceedings of the National Academy of Sciences, 109(20), E1214-E1214.
Goldstein, H., Poole, Jr. C. P. & Safko, J. L. (2001), Classical Mechanics (3rd Edition). Addison-Wesley; 3 edition (June 25, 2001)
Goltsov, A., Faratian, D., Langdon, S. P., Mullen, P., Harrison, D. J., & Bown, J. (2012). Features of the reversible sensitivity-resistance transition in PI3K/PTEN/AKT signalling network after HER2 inhibition. *Cellular signalling*, *24*(2), 493-504.
Greenblatt, M. S., Bennett, W. P., Hollstein, M. & Harris, C. C. (1994). Mutations in The P53 Tumor-Suppressor Gene - Clues To Cancer Etiology And Molecular Pathogenesis. Cancer Research 54(18): 4855-4878
Grover, W. H., Bryan, A. K., Diez-Silva, M., Suresh, S., Higgins, J. M., & Manalis, S. R. (2011). Measuring single-cell density. Proceedings of the National Academy of Sciences, 108(27), 10992-10996.
Guangqi, E., Cao, Y., Bhattacharya, S., Dutta, S., Wang, E., & Mukhopadhyay, D. (2012). Endogenous vascular endothelial growth factor-A (VEGF-A) maintains endothelial cell homeostasis by regulating VEGF receptor-2 transcription. *Journal of Biological Chemistry*, *287*(5), 3029-3041.
Guidolin, D., Albertin, G., & Ribatti, D. (2010). Exploring in vitro angiogenesis by image analysis and mathematical modeling. *Microscopy: science, technology, applications and education*, *2*, 876-884.
Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. cell, 100(1), 57-70.
Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation.

*cell*, *144*(5), 646-674.
Hayashi, M., Majumdar, A., Li, X., Adler, J., Sun, Z., Vertuani, S., Hellberg, C., Mellberg, S., Koch, S., Dimberg A., Koh G. Y., Dejana E., Belting H.G., Affolter M., Thurston G., Holmgren L., Vestweber D. & Claesson-Welsh, L. (2013). VE-PTP regulates VEGFR2 activity in stalk cells to establish endothelial cell polarity and lumen formation. *Nature communications*, *4*, 1672.
http://www.tiem.utk.edu/~gross/bioed/webmodules/celldivision.html
Huang, G. S., Brouwer-Visser, J., Ramirez, M. J., Kim, C. H., Hebert, T. M., Lin, J., ... & Horwitz, S. B. (2010). Insulin-like growth factor 2 expression modulates Taxol resistance and is a candidate biomarker for reduced disease-free survival in ovarian cancer. Clinical Cancer Research, 16(11), 2999-3010.
Illian, J., Penttinen, A., Stoyan, H., & Stoyan, D. (2008). *Statistical analysis and modelling of spatial point patterns* (Vol. 70). John Wiley & Sons.
Indiana University
Bloomington website http://www.informatics.indiana.edu/rocha/complex/csm.html
Jain, R. K. (2005). Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy. Science, 307(5706), 58-62.
Jiujiang Zhu, Simon Coakley, Mike Holcombe, Rod Hose, Rod Smallwood. Cell Adhesion mediates clonal formation of stem and differentiated cell populations (not published)
Johnson, K. L., Kendall, K., & Roberts, A. D. (1971, September). Surface energy and the contact of elastic solids. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 324, No. 1558, pp. 301-313). The Royal Society.
Jones, J. E. (1924, October). On the determination of molecular fields. I. From the variation of the viscosity of a gas with temperature. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 106, No. 738, pp. 441-462). The Royal Society.
Jones, J. E. (1924, October). On the determination of molecular fields. II. From the equation of state of a gas. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 106, No. 738, pp. 463-477). The Royal Society.
Kang, M. H., & Reynolds, C. P. (2009). Bcl-2 inhibitors: targeting mitochondrial apoptotic pathways in cancer therapy. Clinical cancer research, 15(4), 1126-1132.
Kelly, P. N., & Strasser, A. (2011). The role of Bcl-2 and its pro-survival relatives in tumourigenesis and cancer therapy. Cell Death & Differentiation, 18(9), 1414-1424.
Kerr, J. F., Wyllie, A. H., & Currie, A. R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *British journal of cancer*, *26*(4), 239.
King, J. R., & Franks, S. J. (2004). Mathematical analysis of some multi-dimensional tissue-growth models. European Journal of Applied Mathematics, 15(3), 273-295.
Köhn-Luque, A., de Back, W., Starruß, J., Mattiotti, A., Deutsch, A., Pérez-Pomares, J. M., & Herrero, M. A. (2011). Early embryonic vascular patterning by matrix-mediated paracrine signalling: a mathematical model study. *PloS one*, *6*(9), e24175.
Kreeger, P. K., & Lauffenburger, D. A. (2010). Cancer systems biology: a network modeling perspective. *Carcinogenesis*, *31*(1), 2-8.
Kubota, Y., Kleinman, H. K., Martin, G. R., & Lawley, T. J. (1988). Role of laminin and basement membrane in the morphological differentiation of human endothelial cells into capillary-like structures. *The Journal of cell biology*, *107*(4), 1589-1598.

Ladyman, J., Lambert, J., & Wiesner, K. (2013). What is a complex system?. European Journal for Philosophy of Science, 3(1), 33-67.
Lanza, V., Ambrosi, D., & Preziosi, L. (2006). Exogenous control of vascular network formation in vitro: a mathematical model. *Networks and Heterogeneous Media*, *1*(4), 621.
Lee, M. J., Ye, A. S., Gardino, A. K., Heijink, A. M., Sorger, P. K., MacBeath, G., & Yaffe, M. B. (2012). Sequential Application of Anti-Cancer Drugs Enhances Cell Death by Re-wiring Apoptotic Signaling Networks. *Cell*, *149*(4), 780–794. doi:10.1016/j.cell.2012.03.031
Li, P., Zhou, C., Xu, L., & Xiao, H. (2013). Hypoxia enhances stemness of cancer stem cells in glioblastoma: an in vitro study. International journal of medical sciences, 10(4), 399.
Li, T., Kon, N., Jiang, L., Tan, M., Ludwig, T., Zhao, Y., ... & Gu, W. (2012). Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. *Cell*, *149*(6), 1269-1283.
Li, Y. (2012), Parallel Agent-based Cell Model In Vessel Formation. Abertay Research Student Conference
Li, Y., Sampson, A. T., Bown, J. & Deeni, Y. (2013), Understanding tissue morphology: model repurposing using the CoSMoS process. 6th Complex Systems Modelling and Simulation Workshop (the CoSMoS 2013)
Li, Y., Sampson, A. T., Bown, J., Khalil, H. S., & Deeni, Y. (2014). Understanding tissue morphology: model repurposing using the CoSMoS process. *Natural Computing*, 1-20.
Liu, A. M., Wong, K. F., Jiang, X., Qiao, Y., & Luk, J. M. (2012). Regulators of mammalian Hippo pathway in cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, *1826*(2), 357-364.
Liu, A. M., Wong, K. F., Jiang, X., Qiao, Y., & Luk, J. M. (2012). Regulators of mammalian Hippo pathway in cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, *1826*(2), 357-364.
Long, B. L., Rekhi, R., Abrego, A., Jung, J., & Qutub, A. A. (2013). Cells as state machines: cell behavior patterns arise during capillary formation as a function of BDNF and VEGF. *Journal of theoretical biology*, *326*, 43-57.
Luo, J., Solimini, N. L., & Elledge, S. J. (2009). Principles of cancer therapy: oncogene and non-oncogene addiction. Cell, 136(5), 823-837.
Ma, Z., Discher, D. E., & Finkel, T. H. (2012). Mechanical force in T Cell receptor signal initiation. *Frontiers in immunology*, *3*.
Manoussaki, D., Lubkin, S. R., Vemon, R. B., & Murray, J. D. (1996). A mechanical model for the formation of vascular networks in vitro. *Acta biotheoretica*, *44*(3-4), 271-282.
Martin-Belmonte, F., & Perez-Moreno, M. (2012). Epithelial cell polarity, stem cells and cancer. *Nature Reviews Cancer*, *12*(1), 23-38.
Maugis, D. (1992). Adhesion of spheres: the JKR-DMT transition using a Dugdale model. *Journal of Colloid and Interface Science*, *150*(1), 243-269.
Merks, R. M. H., & Koolwijk, P. (2009). Modeling morphogenesis in silico and in vitro: towards quantitative, predictive, cell-based modeling. *Mathematical Modelling of Natural Phenomena*, *4*(04), 149-171.
Merks, R. M., Brodsky, S. V., Goligorksy, M. S., Newman, S. A., & Glazier, J. A. (2006). Cell elongation is key to in silico replication of in vitro vasculogenesis and subsequent remodeling. *Developmental biology*, *289*(1), 44-54.
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E.

(1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, *21*(6), 1087-1092.
Miettinen, M., & Lasota, J. (2006). Gastrointestinal stromal tumors: review on morphology, molecular pathology, prognosis, and differential diagnosis. *Archives of pathology & laboratory medicine*, *130*(10), 1466-1478.
Mogilner, A., Allard, J., & Wollman, R. (2012). Cell polarity: quantitative modeling as a tool in cell biology. *Science*, *336*(6078), 175-179.
Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 17-23.
Mrkvička, T., Muška, M., & Kubečka, J. (2014). Two step estimation for Neyman-Scott point process with inhomogeneous cluster centers. Statistics and Computing, 24(1), 91-100.
Muller, P. A., & Vousden, K. H. (2013). p53 mutations in cancer. Nature cell biology, 15(1), 2-8.
Namy, P., Ohayon, J., & Tracqui, P. (2004). Critical conditions for pattern formation and in vitro tubulogenesis driven by cellular traction fields. *Journal of Theoretical Biology*, *227*(1), 103-120.
Neufeld, G., Cohen, T., Gengrinovitch, S., & Poltorak, Z. (1999). Vascular endothelial growth factor (VEGF) and its receptors. *The FASEB Journal*, *13*(1), 9-22.
Nevins, J. R. (2001). The Rb/E2F pathway and cancer. Human molecular genetics, 10(7), 699-703.
New England Complex Systems Institute (NECSI) website http://necsi.edu/guide/
Newman, T. J. (2005). Modeling multi-cellular systems using sub-cellular elements. *arXiv preprint q-bio/0504028*.
Ortmann, B., Druker, J., & Rocha, S. (2014). Cell cycle progression in response to oxygen levels. *Cellular and Molecular Life Sciences*, *71*(18), 3569-3582.
Ouyang, L., Shi, Z., Zhao, S., Wang, F. T., Zhou, T. T., Liu, B., & Bao, J. K. (2012). Programmed cell death pathways in cancer: a review of apoptosis, autophagy and programmed necrosis. Cell proliferation, 45(6), 487-498.
Ouyang, L., Shi, Z., Zhao, S., Wang, F. T., Zhou, T. T., Liu, B., & Bao, J. K. (2012). Programmed cell death pathways in cancer: a review of apoptosis, autophagy and programmed necrosis. Cell proliferation, 45(6), 487-498.
Painter, K. J., & Hillen, T. (2013). Mathematical modelling of glioma growth: the use of diffusion tensor imaging (DTI) data to predict the anisotropic pathways of cancer invasion. *Journal of theoretical biology*, *323*, 25-39.
Palsson, E. (2007). A 3-D Deformable Ellipsoidal Cell Model with Cell Adhesion and Signaling. In *Single-Cell-Based Models in Biology and Medicine* (pp. 271-299). Birkhäuser Basel.
Palsson, E. (2008). A 3-D model used to explore how cell adhesion and stiffness affect cell sorting and movement in multicellular systems. *Journal of theoretical biology*, *254*(1), 1-13.
Paramonov, L., & Yaliraki, S. N. (2005). The directional contact distance of two ellipsoids: Coarse-grained potentials for anisotropic interactions. *The Journal of chemical physics*, *123*(19), 194111.
Parsa, H., Upadhyay, R., & Sia, S. K. (2011). Uncovering the behaviors of individual cells within a multicellular microvascular community. *Proceedings of the National Academy of Sciences*, *108*(12), 5133-5138.
Passarge, E. (2007). Color atlas of genetics (Vol. 2007). Thieme.
Peng, R., Yao, X., & Ding, J. (2011). Effect of cell anisotropy on differentiation of stem cells on micropatterned surfaces through the controlled single cell adhesion. *Biomaterials*, *32*(32), 8048-8057.

Perram, J. W., & Wertheim, M. S. (1985). Statistical mechanics of hard ellipsoids. I. Overlap algorithm and the contact function. *Journal of Computational Physics*, *58*(3), 409-416.
Perram, J. W., Rasmussen, J., Præstgaard, E., & Lebowitz, J. L. (1996). Ellipsoid contact potential: Theory and relation to overlap potentials. *Physical Review E*, *54*(6), 6565.
Polack, F. A., Andrews, P. S., Ghetiu, T., Read, M., Stepney, S., Timmis, J., & Sampson, A. T. (2010, March). Reflections on the simulation of complex systems for science. In *Engineering of Complex Computer Systems (ICECCS), 2010 15th IEEE International Conference on* (pp. 276-285). IEEE.
Press, W. H. (2007). Numerical recipes 3rd edition: The art of scientific computing. Cambridge university press.
Radszuweit, M., Block, M., Hengstler, J. G., Schöll, E., & Drasdo, D. (2009). Comparing the growth kinetics of cell populations in two and three dimensions. Physical Review E, 79(5), 051907.
Ramis-Conde, I., Chaplain, M. A., Anderson, A. R., & Drasdo, D. (2009). Multi-scale modelling of cancer cell intravasation: the role of cadherins in metastasis. *Physical biology*, *6*(1), 016008.
Ramis-Conde, I., Chaplain, M. A., Anderson, A. R., & Drasdo, D. (2009). Multi-scale modelling of cancer cell intravasation: the role of cadherins in metastasis. *Physical biology*, *6*(1), 016008.
Ramis-Conde, I., Drasdo, D., Anderson, A. R., & Chaplain, M. A. (2008). Modeling the influence of the E-cadherin-β-catenin pathway in cancer cell invasion: a multiscale approach. *Biophysical journal*, *95*(1), 155-165.
Ramis-Conde, I., Drasdo, D., Anderson, A. R., & Chaplain, M. A. (2008). Modeling the influence of the E-cadherin-β-catenin pathway in cancer cell invasion: a multiscale approach. *Biophysical journal*, *95*(1), 155-165.
Rasband, M. N. (2010). The axon initial segment and the maintenance of neuronal polarity. *Nature Reviews Neuroscience*, *11*(8), 552-562.
Reinders, J. (2007). *Intel threading building blocks: outfitting C++ for multi-core processor parallelism*. " O'Reilly Media, Inc.".
Rieu, J. P., Upadhyaya, A., Glazier, J. A., Ouchi, N. B., & Sawada, Y. (2000). Diffusion and deformations of single hydra cells in cellular aggregates. *Biophysical Journal*, *79*(4), 1903-1914.
Sadava, D., Hillis, D., Heller, C., & Berenbaum, M. (2011). Life: The science of biology. (9th ed.) Sunderland, MA: Sinauer Associates Inc.
Samani, A. A., Yakar, S., LeRoith, D., & Brodt, P. (2007). The role of the IGF system in cancer growth and metastasis: overview and recent insights. Endocrine reviews, 28(1), 20-47.
Scianna, M., & Preziosi, L. (2013). *Cellular Potts models: multiscale extensions and biological applications*. CRC Press.
Scianna, M., Bell, C. G., & Preziosi, L. (2013). A review of mathematical models for the formation of vascular networks. *Journal of theoretical biology*, *333*, 174-209.
Serini, G., Ambrosi, D., Giraudo, E., Gamba, A., Preziosi, L., & Bussolino, F. (2003). Modeling the early stages of vascular network assembly. *The EMBO Journal*, *22*(8), 1771-1779.
Shirinifard, A., Gens, J. S., Zaitlen, B. L., Popławski, N. J., Swat, M., & Glazier, J. A. (2009). 3D multi-cell simulation of tumor growth and angiogenesis. *PLoS One*, *4*(10), e7190.
Stauffer, D., & Aharony, A. (1994). Introduction to percolation theory. CRC press.

Steinberg, M. S. (1962). Mechanism of tissue reconstruction by dissociated cells, II: Time-course of events. *Science*, *137*(3532), 762-763.
Steinberg, M. S. (1962). On the mechanism of tissue reconstruction by dissociated cells, I. Population kinetics, differential adhesiveness, and the absence of directed migration. *Proceedings of the National Academy of Sciences of the United States of America*, *48*(9), 1577.
Steinberg, M. S. (1962). On the mechanism of tissue reconstruction by dissociated cells, III. Free energy relations and the reorganization of fused, heteronomic tissue fragments. *Proceedings of the National Academy of Sciences of the United States of America*, *48*(10), 1769.
Stolarska, M. A., Kim, Y., & Othmer, H. G. (2009). Multi-scale models of cell and tissue dynamics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *367*(1902), 3525-3553.
Sundfeldt, K. (2003). Cell–cell adhesion in the normal ovary and ovarian tumors of epithelial origin; an exception to the rule. *Molecular and cellular endocrinology*, *202*(1), 89-96.
Thomas, S., Quinn, B. A., Das, S. K., Dash, R., Emdad, L., Dasgupta, S., ... & Fisher, P. B. (2013). Targeting the Bcl-2 family for cancer therapy. Expert opinion on therapeutic targets, 17(1), 61-75.
Toettcher, J. E., Loewer, A., Ostheimer, G. J., Yaffe, M. B., Tidor, B., & Lahav, G. (2009). Distinct mechanisms act in concert to mediate cell cycle arrest. Proceedings of the National Academy of Sciences, 106(3), 785-790.
Treloar, K. K., Simpson, M. J., Binder, B. J., McElwain, D. S., & Baker, R. E. (2014). Assessing the role of spatial correlations during collective cell spreading. Scientific reports, 4.
Vaudry, D., Stork, P. J. S., Lazarovici, P., & Eiden, L. E. (2002). Signaling pathways for PC12 cell differentiation: making the right connections. *Science*, *296*(5573), 1648-1649.
Vempati, P., Popel, A. S., & Mac Gabhann, F. (2011). Formation of VEGF isoform-specific spatial distributions governing angiogenesis: computational analysis. *BMC systems biology*, *5*(1), 59.
Vempati, P., Popel, A. S., & Mac Gabhann, F. (2014). Extracellular regulation of VEGF: isoforms, proteolysis, and vascular patterning. *Cytokine & growth factor reviews*, *25*(1), 1-19.
Vousden, K. H., & Prives, C. (2009). Blinded by the light: the growing complexity of p53. Cell, 137(3), 413-431.